The Dissertation Committee for Alexandra Nikolaos Charbi
certifies that this is the approved version of the following dissertation:

Quality Reporting, Bonus Payments and Welfare in Medicare Advantage

Committee:

Robert Town, Supervisor

Daniel Ackerberg

Eugenio Miravete

Marika Cabral

Naoki Aizawa

Quality Reporting, Bonus Payments and Welfare in Medicare Advantage

by

Alexandra Nikolaos **Charbi**

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2021

Dedicated to Nick, Stella, and Alexandra.

Acknowledgments

I wish to thank the multitude of people who supported me during this journey. First and foremost, I would like to thank my parents Nick and Stella, and my beloved grandmother Alexandra. They believed in me, they invested in me in thousands of different ways, and they supported every path I took. Two of them are not with me anymore, but I am sure that Stella is as proud as they would be.

I would like to thank Nikos Vettas and Yannis Katsoulacos who encouraged me to continue my studies in the U.S. and to broaden my horizons. I am also deeply grateful to David Miller and Tilman Börgers who supported my second chance and Tom Wiseman who believed that I deserved it!

Words cannot describe how grateful I am for all the guidance, support, and encouragement my advisor Robert Town provided me with. He invested many hours in me and my project and his advice was helpful at all levels: research, professional and personal. Daniel Ackerberg also provided me with unique feedback and advice. I am grateful for the many hours I spent in his office learning how to put structure on my thoughts and how to use multiple econometric tools. In the end, he convinced me, as I know he wishes for every student, that I can estimate everything I want! I further wish to thank Eugenio Miravete who challenged me in multiple ways and made me stronger along this path. I am also grateful to Marika Cabral and all the feedback she provided me with. Simply put, I could not ask for a better committee!!! Last but not least, I would like to kindly thank Micheal Geruso and Jorge Balat for the extended conversations we have had all these years.

I am also grateful to the unique friends I gained throughout this process, Laurien Gilbert, Christina Zafeiridou, Steven Hamilton, Christopher Sullivan and Austin Bean. To Chara Tzanetaki and Philip Atzemoglou with whom I have been sharing a great friendship for the last 16 years. All my friends believed in me and stood by me in every possible way. Our extended conversations and video chats helped me move forward. Further, I would like to thank my classmates in the IO writing seminar for the great exchange of thoughts we've had. And last but not least, Houndstooth Coffee for the many hours that hosted me and the good coffee that provided me across its different locations in Austin!

Quality Reporting, Bonus Payments and Welfare in Medicare Advantage

Publication No. ________________

Alexandra Nikolaos Charbi, Ph.D.
The University of Texas at Austin, 2021

Supervisor: Robert Town

When consumers are imperfectly informed about the quality of a product, market forces do not incentivize firms to provide the socially optimal level of quality. Imperfect information is a recognized and frequent market failure in the context of public health and has led to initiatives aimed at increasing consumers' access to information and at incentivizing firms to provide higher quality services. This study analyzes the welfare effects of quality disclosure and quality subsidies on the Medicare Advantage (MA) market. MA is a subsidized program that provides health insurance to the elderly and disabled population in the U.S. as an alternative to Traditional Medicare.

The study begins by introducing the institutional background of the market that provides a unique setup to analyze. On the demand side, consumers receive information on the quality of health insurance plans through a

Star Rating System (SRS). On the supply side, higher-rated insurers receive a quality-linked subsidy through a Quality Bonus Program (QBP).

The second chapter provides evidence that consumers do not respond to the information they receive. The question then arises, "Why is this happening?" Is is because consumers are not aware of the SRS? Is it because consumers do not value the information they receive through the SRS? Or, is it because consumers do not care about the dimension of quality the SRS is informing them on? At the same time, star ratings increase over time. Simple density distribution graphs suggest that this increase is motivated by the financial incentives that QBP provides insurers with. Those two observations raise the question of what the relative impacts of the two policies together are on welfare.

The third chapter describes a survey I designed and conducted to answer the first set of questions regarding the demand side of the market–Do consumer know about the SRS? Do consumers value the information they receive? Do consumers about the dimension of quality the star ratings are informing them on?–Surveying Medicare-eligible individuals, I find that 80% of the population is *unaware* of the SRS. In the survey, I also conduct a conjoint analysis to elicit preferences for star ratings. I find that respondents who reported they were *aware* of the SRS place a monthly value of $25 on an extra star rating; slightly more than the ones who reported they were *unaware*.

The fourth chapter presents a structural equilibrium model of supply and demand that separately identifies and quantifies the relative impacts of each policy on welfare. The model also incorporates the survey results by flexibly allowing for different consumer types: those who are *unaware* and those who do *not care* about the SRS. I combine the survey stated preference with revealed preference choice data and estimate a Bayesian learning discrete choice model. On the supply side, insurers endogenously choose price and quality. My analysis shows that although both the SRS and the QBP lead to higher quality, welfare improvement is very small compared to the incurred costs. In particular, 75% of the expenditures spent on the QBP is not rationalized by any welfare improvement.

The final chapter concludes and states potential paths for future research.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

When consumers are imperfectly informed about the quality of a product, market forces are thought not to provide the correct incentives for firms to invest in the socially optimal level of quality. This has long been known and dates back to [8], [7], [86]. This observation has led many public and private initiatives to provide consumers with better product quality information, including Consumer Reports, Yelp reviews, health departments' restaurant hygiene scores, and school performance information. However, the provision of product quality information may provide firms insufficient incentives (from a welfare perspective) for quality investments. This may be because of remaining information frictions, the inability of performance measures to capture the relevant quality domains, or firm market power. When this is true, the government can intervene and provide direct incentives for firms to invest in product quality. They can do this through testing (e.g., pharmaceuticals), licensure (e.g., physicians), or, as this study analyzes, through direct quality subsidies.

In my thesis, I study the Medicare Advantage (MA) Star Rating System (SRS). MA is a program that allows Medicare beneficiaries to bypass the Traditional Medicare (TM) program, which provides health insurance to the elderly and disabled population in the U.S., and enroll in a private plan that is responsible for the care they receive. The plans receive a subsidy payment for each of their enrollees. The SRS measures the performance of these MA plans using more than 40 measures and then maps the outcomes of these measures into a single-dimensional star rating between 1 and 5 in half-star increments for each plan. The SRS was introduced in 2008. Beginning in 2012, the Centers for Medicare and Medicaid Services (CMS), the agency that oversees the Medicare program, instituted a policy, named Quality Bonus Program (QBP), that linked the subsidies the MA plans receive to their star rating. In my dissertation, I study the impact of the SRS and the QBP on consumer demand, plan pricing and quality, and ultimately, welfare.

Although it is generally accepted that both policies can improve quality, and probably welfare, they can also entail significant costs: costs that refer to spending for generating the relevant information and its delivery to consumers, and costs for the bonuses insurers receive. Then, the arising question is whether the incurred cost is justified by the welfare improvement the market may experience. The answer depends on both the "demand-and supply-side" environment of the market. On the demand side, three factors are critical. First, to what extent are consumers aware of the information they receive? Second, to what extent do consumers value the information they receive? Third, to what extent do consumers have significant preferences for the part of quality the SRS is informing them on? On the supply side, it is critical to understand how insurers are incentivized to produce and price quality given the demand structure and the financial incentives they receive. For example, when consumers are not aware of the available information, insurers do not have strong incentives to produce high quality plans. Even worse, they can impose high prices on low-quality plans taking advantage of consumers' inaccurate beliefs about quality. Similarly, when consumers do not value quality, insurers do not have any incentive to invest in it, unless they receive extra bonuses for doing so. In this scenario, although quality increases, consumer surplus remains unaffected leaving the cost of the bonuses unjustified.

I begin my analysis by providing evidence that consumers do not care about star ratings, but at the same time they are enrolled in high quality plans. This is not very surprising as health insurance markets are notorious for the high levels of inattention they are plagued by. Also, the data show that the market experiences significant quality improvement over time that is mainly due to the financial incentives insurers receive through the QBP. These two observations together raise the main question of this study: "What are the relative impacts of the SRS and the QBP on welfare?" Knowing to what extent the combination of the implemented policies can lead to better outcomes and to what extent the cost that is incurred is justified by welfare improvement is important from an economic perspective. Furthermore, knowing what combination of policies works better can lead to better policy recommendations.

I build a full demand and supply equilibrium model to study whether the combination of the implemented policies improves welfare and to what extent the obtained cost is justified. On the demand side, I investigate how consumers form beliefs about quality and how much they value the information that is provided to them. In doing so, I build a Bayesian learning discrete choice model in the spirit of [26], which also introduces a type of information friction that refers to consumer *awareness* with respect to the SRS. On the supply side, I investigate whether insurers under- or over-provide quality in building a model in which firms choose prices and quality while they endogenize the information environment and the financial incentives they receive.

Introducing a type of friction that refers to *awareness* is a novel part of this study. Although previous research has acknowledged the importance of information awareness, to the best of my knowledge data limitations have prevented empirical investigation of this friction before this study. In aggregate or individual choice-level data, consumers who do not care about a product characteristic are observationally equal to consumers who do not know about it. Usually, the literature assumes perfect information, fully attributing consumer choices to preferences. However, it is possible that, if consumers had known about this product characteristic, they would care about it. In such a case, the assumption of perfect information leads to wrong welfare implications. Distinguishing these consumer types is not always necessary, but in a study such as mine that analyzes two policies that heavily depend on the information environment it is critical for welfare predictions. To the best of my knowledge, I am the first to empirically investigate this type of information friction.[1]

I model consumer preferences for quality with a random coefficient that is governed by a normal distribution with a mass point at zero, and I allow consumers to make their choices according to their preferences, as in [13]. The challenge arises in distinguishing consumers who are not *aware* of the SRS from consumers who do not *care* about quality. To overcome this challenge, I design and run an electronic survey that yields unique data on 624 nationally representative Medicare-eligible individuals. In the survey, I recover the proportion of consumers who were aware/unaware of the SRS, and I further elicit preferences for star ratings. To recover the proportion of consumers who were aware of the SRS, I directly ask respondents whether they knew anything about it. To elicit preferences for star ratings, I conduct a conjoint experiment in the spirit of [11], providing respondents a series of choice sets to choose from with a trade-off between prices and star ratings. I find that 77.4% of the Medicare population is unaware of the SRS. I estimate respondent preferences using the method of Simulated Maximum Likelihood. Results from the conjoint analysis show that 17% of the initially aware respondents did not care about the star ratings, and 25% of the newly aware respondents did not care about star ratings. Overall, initially aware respondents assign a $25 monthly value per star as compared to respondents who were newly aware,

[1] More recently, [2] develop a method to estimate discrete choice models in which, when the assumptions of a canonical model are combined with a set of mild restrictions, the use of choice data only is sufficient to identify consumer preferences when choices are not fully informed. Their model has not yet been applied empirically.

who assigned a $20 monthly value per star.

I proceed by estimating the main Bayesian learning demand model. In my setup, consumers have some prior beliefs about MA plan quality before the introduction of the SRS. After the introduction of the SRS, consumers that are aware of it, update their beliefs based on the signals the star ratings give them. If consumers are not aware of the SRS, they behave as if they were in the pre-SRS period based on their prior beliefs. I impose the consumer preferences for star ratings I recovered in the survey on my main demand model to estimate the rest of the model parameters.

My estimation follows [76] and the algorithm used by [14]. The key point of estimation exploits a population moment condition that requires a set of exogenous instrumental variables to form a non-linear Generalized Method of Moments estimator. Combining stated preference choice survey data with revealed preference choice aggregate data requires me to control for possible differences in the idiosyncratic tastes of the survey respondents and the actual consumers. In doing so, I use a rescale parameter which also has the advantage of preserving the shape of the preference distribution that I recovered from the survey in the main model. The estimates predict that consumers are moderately elastic with respect to price, with own price elasticity close to -1.16. The most valuable benefit for consumers is drug coverage; consumers are willing to pay up to $170/month for drug coverage, which is six times more than what they are willing to pay for an extra star rating.

Once I estimate the demand parameters, I use them to separately estimate the "supply-side" parameters. In my model, insurers choose both prices and quality, endogenizing the information environment and the financial incentives they receive from Medicare. I utilize the competitive environment of the market, and after I invert the first-order conditions for prices, I recover insurers' implied marginal cost of the services they provide. I assume that insurers cannot perfectly predict the final level of star rating that will arise, since there is a long list of metrics that determine the final outcome. I generate a continuous measure of quality, given CMS's provided algorithms. Then, I utilize the considerable variation I observe around the thresholds at which a new star rating is assigned, and estimate via the Simulated Maximum Likelihood method the parameter that governs insurers' uncertainty when they decide the quality level they will produce. I find that an extra star rating costs insurers slightly less than $25/month and that the final level of quality that results is noisy with a standard deviation of 1.4 stars. Finally, I provide evidence that the model I estimate fits the data remarkably well.

In my counterfactual analyses, I quantify the relative impacts of the SRS and the QBP on firm profitability, consumer surplus, and government expenditures. I compare outcomes that arise under three different levels of information structure: (i) full information, (ii) current level of information, and (iii) no information, and two supply-side regulatory environments: (i) with and (ii) without quality subsidies. Not surprisingly, I find that both the SRS and QBP improve quality. As consumers become more aware of the existing information, plan quality goes up. Similarly, when insurers are provided financial incentives to improve quality, they do so.

Interestingly, although both policies improve quality significantly, consumer surplus only increases modestly. This pattern is observed either because consumers do not value the dimension of quality the star ratings represent or because star ratings are not good signals of quality. On average, firms become less profitable as consumers become more informed about the SRS. This is not surprising. When consumers are not aware of the true plan quality, firms impose high prices on low-quality plans and thus become profitable.

Further, I find that the same level of quality that is observed in practice through the combination of "demand-and supply-side" policies can be achieved through the "demand-side" policy alone if the informed share of the population increased from 22% to 50%. This result suggests that the government may be able to avoid the substantial costs of the QBP, which amount to $1.7b per year, if it can find a more cost effective way of informing consumers about plan quality.

The striking finding of my analysis is that the cost that is incurred to increase quality via bonus subsidies is not justified by the total welfare improvement. Of the $1.7b Medicare spends annually on average to improve welfare via the QBP, only 25% is rationalized by the increased aggregate surplus that flows to consumers and producers. My analysis defines welfare traditionally as the sum of consumer surplus and firm profitability without taking into account any other benefits that might be realized due to higher quality. Hence, it remains unclear whether the part of the Medicare expenditures that is not rationalized generates a positive externality or whether it is just lost. However, my analysis suggests that, for a star rating to justify the extra spending, it should generate a positive externality close to the amount of $6b annually.

The remainder of the study proceeds as follows. Section 1.1 presents the institutional background on the MA market and the policies of interest. Section 1.2 describes the literature to which this paper is related. Chapter 2 describes the data and provides a preliminary analysis of the first major observations. Chapter 3 describes the survey I conducted and analyzes in detail the results that arose. Chapter 4 describes the empirical framework for the demand and the supply side model along with the results and the welfare analysis. Chapter 5 concludes the study.

1.1 Institutional Background

Medicare, enacted in 1965, is a public program that provides health insurance to the elderly and disabled population in the U.S. regardless of income or medical history.[2] It is the largest health insurance program in the country, and as of 2020 it comprised 3% of the Gross Domestic Product. It is currently managed by the CMS and covers hospital care (Part A) and medical care (Part B). Since 2006, it has also been providing prescription drug coverage (Part D). Under its fee-for-service structure, also known as TM, providers that choose to participate in the program are required to treat beneficiaries

[2]Disabilities mainly refer to End Stage Renal Disease and Amyotropic Lateral Sclerosis.

in exchange for a reimbursement. Beneficiaries pay a fixed annual amount for Parts A and B as well as fixed copays, coinsurance rates, and deductibles, depending on the medical care they receive.

1.1.1 The Medicare Advantage program

MA allows beneficiaries to opt out of traditional fee-for-service Medicare and enroll in a private insurance plan. The program, originally called Medicare+Choice (Part C), originates in the Tax Equity and Fiscal Responsibility Act of 1982, which targeted the expansion of plan options and cost reduction as a result of the managed care competition. MA plans must provide at least Part A and B coverage; sometimes they also offer Part D coverage. Beneficiaries choosing an MA plan receive medical benefits from their plan exclusively. Although at first the program did not receive a lot of attention, its popularity has increased during the last twenty years. In 2019, 95% of Medicare beneficiaries had the option of choosing an MA plan in their county, and MA plans served 33% of the Medicare population.[3]

After MA insurers have proven they meet a series of financial and administrative criteria, they sign contracts with Medicare for the set of plans they will offer to different counties.[4] A contract refers to a specific product type, such as a Health Maintenance Organization (HMO), Preferred Provider Organization (PPO), and Private Fee For Service (PFFS), that covers specific

[3]See [71] for details regarding the history of MA.

[4]They need to prove that they can bear the financial cost of offering plans. They are also required to build the necessary infrastructure for their plans to operate efficiently.

geographical areas and includes a set of plans that all share the same network of providers.[5] A plan refers to a specific combination of premium, copayment, coinsurance rate, deductible, and other benefits, such as drug, vision, dental, and hearing. Each insurer is required to provide at least one plan offering Part D benefits; these plans are called MA-PD plans.

Insurers that choose to participate in the program have to fully accept the risk patients carry.[6] In exchange for the services insurers offer, they receive a monthly payment per patient. This payment depends on risk adjusted benchmark (capitation) rates that Medicare determines depending on the per capita fee-for-service spending within a county.[7] Every year after the release of the benchmark rates, insurers submit a “bid” for each plan they offer depending on their cost. If the submitted “bid” is above the county benchmark, the insurer receives an amount equal to the benchmark, and to cover the difference, adds extra charges to the standard Part B and/or to any supplemental premiums. If the submitted “bid” is below the county benchmark,

[5]HMOs deliver care through providers who work directly for the insurance firm. Beneficiaries who buy these plans are required to choose their primary provider, and they must have a referral to visit a specialist. HMO enrollees may see a non-network provider, but they will have to pay extra. HMOs also have restrictions on the number of visits, tests, or treatments their enrollees can receive. PPOs are similar to HMOs, but they are more flexible in terms of networks, patients' ability to see specialists, plan costs, and coverage for out-of-network services. Finally, PFFS plans resemble TM plans in access, benefits, and reimbursement to providers; PFFS offers a flat reimbursement rate per procedure, and providers can choose on a case-by-case or service-by-service basis whether or not to accept patients.

[6]Risk refers to the risk beneficiaries carry associated with their heath status.

[7]For different plans an insurer offers, the insurer can receive different levels of reimbursement.

the insurer receives an amount equal to the "bid" and an additional "rebate" (as a reward), the amount of which is a fixed percentage share of the difference between the "bid" and the benchmark.[8] Insurers are required to pass the actuarial value of the "rebate" to enrollees through additional benefits or by reducing the standard Part B premium. Consequently, insurers compete with each other on premiums, and more recently they have started competing on dimensions of quality, such as benefit design, size of network and advertising. ([29], [6], [73])

Medicare beneficiaries are automatically signed up for Medicare if they receive Social Security or Railroad Retirement Board benefits or submit an application by calling the Medicare line. Every year during the Open Enrollment period (October, 15 – December, 7), beneficiaries are offered a variety of plans in the county they live. To find information about the available plans, they can visit Medicare's Plan Finder website (`https://www.medicare.gov`) or read the CMS "Medicare & You" handbook. Additionally, they are provided with information by the insurers' advertisements. Finally, they have the option to consult with counselors from the State Health Insurance Program and/or independent brokers. Within this period and according to their needs, beneficiaries choose to either enroll in an MA plan or receive coverage from TM.[9]

[8]At the beginning of my data, the "rebate" was equal to 75% of the difference between the "bid" and the benchmark. After the introduction of the QBP program, this fixed percentage changed depending on the star ratings each plan received. See more details in section 1.1.3.

[9]In case they first become eligible for Medicare during a different period or a major life

From a beneficiary's point of view, the decision between TM and MA involves a trade-off of higher out-of-pocket cost in return for less restrictive coverage. Typically, TM enrollees face high out-of-pocket costs, which they cover by buying supplemental Medigap policies. In contrast, MA plans offer more generous cost sharing and may also provide extra benefits, such as vision, dental, hearing, and drug coverage. However unlike TM, MA plans impose limited access to providers. Most MA beneficiaries are enrolled in either HMO or PPO plans, with various restrictions on the provider networks.

1.1.2 The Star Rating System

CMS began constructing and publishing MA and Medicare Part D star ratings in 2008. The overarching goal of the star rating initiative is to provide beneficiaries information on plan quality to facilitate better plan choice. The ratings are presented on a scale from 1 to 5 in half-star increments.

The rules governing the assignment of a final star rating to an insurer are complicated and have changed over time.[10] Broadly speaking, a variety of data sources (coming from the Consumer Assessment of Healthcare Providers and Systems (CAHPS) and the Healthcare Effectiveness Data and Information Set (HEDIS) surveys) is used to collect information on a set of 30 to 44 performance measures spanning 5 to 9 categories, which are aligned with the

change happens, they have the option to enroll at any time of the year. More recently, they also have the option to enroll in a 5-star plan at any time of the year.

[10]Section 1.1 in the appendix analyzes in detail the main rules that govern the assignment of star ratings to insurers.

government's goals. These measures are non-financial and refer to some of the clinical outcomes of the plans' enrollees, the way plans help their enrollees manage their chronic conditions, member experience with the plan, access to medical care, and customer service. CMS uses relative distribution and clustering techniques to convert the original scores of these measures into star ratings. Next, it uses a weighting system to summarize the individual star ratings into an average "summary rate" that is rounded to the nearest half star.[11] Table 1.1 shows the exact measures along with the weights CMS used in 2015. Star ratings are assigned at a contract level, and every plan under the same contract gets the same star rating. Contracts receiving less than two star ratings are excluded from the market until they improve their quality.

Every year, before the enrollment period begins, CMS officially discloses star ratings along with other plan characteristics on Medicare's Plan Finder website. Beneficiaries can visit Plan Finder to compare plans in different dimensions. Figure 1.1 provides an example of the plan characteristics beneficiaries see and compare on Plan Finder. It is important to notice in this example that as star ratings reflect a dimension of quality that refers to customer satisfaction and general clinical outcomes, they are distinguished from the general benefit design of a plan. Other sources where beneficiaries can find information about star ratings are insurers' websites and/or their mail-

[11]For example, a plan with a 3.75 summary rate is rounded up to a 4-star rating, while a plan with a 3.74 summary rate is rounded down to a 3.5-star rating. For MA plans offering Part D benefits (MA-PD), Part C and D "summary rates" are combined to create an "overall rate". For MA plans not offering Part D benefits, the "summary rate" is also the "overall rate".

ings. However, it is important to mention that insurers are not mandated to disclose star rating information. Even when they do so, it is often the case that consumers put costly effort to find the relevant information compared to the effort they put to find information with respect to other plan characteristics. Lastly, although beneficiaries can generally find information about plan characteristics in the "Medicare" handbook, that is distributed to them annually, CMS chooses not to include star ratings information in the book. All these observations together imply that consumers may sign up for plans without being aware of star ratings.

Criticisms on the SRS: The SRS has been criticized for its use both on the demand side and the supply side. On the demand side, the main concerns refer to its use as a guide to consumers choosing among MA plans. As mentioned, a major concern refers to the possibility that consumers choose MA plans without knowing anything about the star ratings information. In addition, the complexity of the SRS and the abundance of information a final star rating conveys make it difficult for consumers to understand and process it so that they can use it effectively in their choice decision. On the supply side, the main concern refers to the possibility that insurers "manipulate" the SRS as a quality metric system. Although CMS has been changing the rules it follows to assign the final star ratings,[12] insurers are aware of the main algorithm CMS follows along with the metrics that are heavily weighted. Hence, insur-

[12]These changes refer to the weights each metric is assigned with, the specific measures it uses, etc. Sometimes these changes are pre-announced, whereas other times they are sudden.

ers can either invest in quality targeting metrics that are "easier" for them to increase or they can "game" the system showing that they are of higher quality when their actual quality does not change. In this study, I focus on the possibility that consumers might be *"unaware"* of the SRS. Further, my analysis assumes that insurers target an overall star rating instead of specific metrics and excludes the possibility of star ratings manipulation and leaves it for future analysis. These concerns are left for future analysis.

1.1.3 The Quality Bonus Program

The QBP originates in the amendments of the 2010 Patient Protection and Affordable Care Act, which targeted quality improvement of health insurance plans. It is the first Pay-For-Performance program in the market for health insurance and provides insurers different levels of bonuses depending on the star rating they receive.[13] In March 2010, CMS announced that as of the beginning of 2012 contracts receiving 4 or more star ratings would receive a 5% higher benchmark rate than all other contracts. To further incentivize more insurers to increase quality, in November 2010, Medicare decided to begin the program with a "demonstration project". This "demonstration project" would last for three years (2012 – 2014) and would expand bonuses to plans receiving lower star ratings. In 2015, the original bonus scheme plan went into

[13] Pay-For-Performance programs are not new. They have been used in many industries, and recently they have been used in the healthcare sector to incentivize providers to improve their quality.

effect.[14] Table 1.2 shows how the bonus schemes of the program evolved over time.[15]

Bonus payments are paid per enrollee and are calculated as a share of the MA benchmarks, which vary by county. Lastly, bonuses for new contracts are 3.5% of the benchmark, and plans failing to report their quality do not receive any bonus payments.

1.1.4 Summary of the payment structure after the implementation of SRS and QBP

The period between the data collection and the final release of star ratings lasts two years. During this period, providers deliver healthcare services, CAHPS and HEDIS data are released, and CMS collects these data and calculates the final star ratings. In the meantime, insurers choose their prices and the star ratings they will target each year, consumers choose the plans that match their needs, and CMS pays insurers accordingly. The following example describes this sequence of events covering a two-year period (January 2019 – December 2020):

1. From January 2019 to August 2019, insurers deliver healthcare services to their enrollees.

[14]The demonstration project also allowed bonuses for contracts in some counties with special demographic characteristics to be double.

[15]Since the implementation of the QBP, the rebates, as defined by a "fixed" percentage of the difference between the benchmark rates and the "bids", have also changed. Table 1.3 shows exactly how they have evolved over time.

2. Based on the insurer performance during this period, CAHPS and HEDIS measures are released in September 2019.

3. From January 2020 to September 2020, CMS collects these data and calculates the star ratings for the enrollment period of 2021. During this period, it also announces changes that may occur regarding the calculation of the star ratings in the future.

4. In October 2020, the star ratings are released for the plans that will serve the market in 2021.

5. After the release of the star ratings, the benchmark rates will be finalized and insurers will submit their bids.

6. The official enrollment period is between October, 15th – December, 7th 2020. During this period, consumers see the plans offered in the area they live, observe prices, star ratings and other plan characteristics, and decide which plan they will enroll in.

7. The final CMS payment is given to every insurer.

Figure 1.2 also shows a graphical representation of the above example. In summary, the final payment from CMS to an MA insurer for a plan j he offers in market m and year t is determined as follows:

$$Payment_{jmt} = \begin{cases} B_{mt} \cdot \overline{RA}_{jt} \cdot \psi_{jt} & \text{if } bid_{jt} \geq B_{mt} \\ (bid_{jt} + \zeta_{jt} \underbrace{(B_{mt} - bid_{jt})}_{rebate_{kjt}}) \cdot \overline{RA}_{jt} \cdot \psi_{jt} & \text{if } bid_{jt} < B_{mt} \end{cases}, \tag{1.1}$$

where B_{mt} represents the benchmark rates, $\overline{RA}_{jt}$ represents the average risk adjustment, ψ_{jt} is the bonus plans receive depending on the star rating they receive, and ζ_{jt} is the fixed percentage rate of the rebate as it has evolved over time.[16] In case a plan also offers Part D benefits, the final payment will be a sum of the above plus an additional amount that results from a similar bidding system. The final premium of a plan is determined as follows:

$$p_{jmt} = \begin{cases} bid_{jt} - Payment_{jmt} & \text{if } bid_{jt} \geq B_{mt} \\ 0 & \text{otherwise} \end{cases}, \tag{1.2}$$

and for plans that also offer Part D coverage an additional premium is added.

Criticisms on the QBP: The main concern around the QBP has always been its cost. The first three years of the program increased the size of the bonus payments as well as the number of contracts receiving them, providing bonuses to the vast majority of MA contracts with total spending approaching $8.35b. Part of this cost is rationalized by the consumer willingness to pay for high-star plans. However, if consumers are not aware of the existence of the SRS or if they are not willing to pay for high-star plans, then the question that arises is whether there are more efficient ways through which the same quality outcome can be realized at a lower cost.

[16]This fixed percentage rate ζ_{jt} varied over time. Before 2012 it was fixed at 75%, between 2012 and 2014 it was a blend of the pre- and post-that period rebate amounts, and after 2014 it started varying depending on plans' star ratings; for plans receiving $4-4.5$ stars it was 70%, for plans receiving $3.5-4$ stars it was 65%, and for the rest it was 50%. Table 1.3 shows exactly how it varied over time. Also, the $(1-\zeta_{jt})$% was obtained by the Trust Funds as savings.

1.2 Literature Review

My study closely relates to papers that analyze the welfare effects of information provision on quality through reporting cards, star ratings ([9], [94], [55], [56], [30], [63]), or advertising ([5], [82]). These papers assume that information fully reaches consumers and, consequently, estimate the effect of the information provision's intent to treat with information provision. I allow for the possibility that information may not fully reach consumers. I identify the percent of consumers who are aware of the available information and quantify the effect of the treatment of information provision on the treated consumers.

My study extends the literature that analyzes different types of information frictions that result in imperfectly informed decision-makers. This literature goes back to [8], who points out that imperfect information is a main friction that prevents consumers from making efficient choice decisions.[17] A common type of information friction that researchers have focused on refers to rational inattention where consumers decide not to avail themselves of the information provided because of the cost of obtaining that information ([87], [66], [21], [58], [59]). Another form of rational inattention refers to switching or searching costs ([77], [51], [43] , [79], [97]). Other researchers have focused on cases where consumers may not be able to comprehend information ([15],

[17]For an extended literature review on the quality of consumer decision-making in the market for health insurance, see [60].

[16]). Seminal papers in this category include [3], [4], [54], [53].[18] I focus on a type of friction that refers to *awareness* and is simply described by a situation in which consumers are not aware that the information is available. [2] theoretically suggest a way to identify *awareness* levels using revealed preference choice data. To the best of my knowledge, I am the first to empirically address this friction and analyze the welfare outcomes that arise in a fully informed market.

Closer to my study, [17] is concerned with potential limited price transparency. While in his analysis he endogenizes the decisions of consumers to become aware of the information, I treat *awareness* as an exogenous variable. Methodologically, in his study, he uses website traffic log data to identify the percent of the population that receives information. In contrast, I utilize survey data; website (PlanFinder) traffic log data would lead to biased results, as a high percent of the website users are also researchers.

My study is related to a large body of literature that analyzes the Medicare Advantage program. Some studies in the MA literature have focused on the pass-through rates of the MA subsidies to the final consumers ([85], [38], [20]), while others have focused on adverse selection and its welfare implications ([64]). While most of these studies have focused on financial aspects of quality, this paper focuses on non-financial aspects that are also welfare relevant. Intuitively, my study is close to [6], who analyze the impact of in-

[18]For a full list of the literature, see [50].

formation provision in the form of advertising. Methodologically, it is close to [91] and [29], who build structural models for welfare analysis.

I significantly contribute to the part of the literature that analyzes the SRA and the QBP. Existing studies have analyzed the two programs separately from each other. [32] study the effects of star ratings on enrollment, finding conflicting evidence across the different years of their analysis.[19] [62] study the effect of the bonus schemes on plan quality. My analysis is closer to that of [89], who builds a structural model to analyze the equilibrium effects of the "demonstration project".[20] To the best of my knowledge, I am the first researcher to analyze both programs, quantifying the relative impact of each one on welfare. Importantly, I treat star ratings as noisy signals of quality; not as quality per se. On the demand side, I identify consumers who might be unaware of the Star Rating System by surveying Medicare-eligible individuals. I further conduct a conjoint experiment to quantify how much consumers value star ratings, and I combine stated preference with revealed preference choice data to estimate my demand model. Using survey data helps me assess both consumers' information set and preferences, as well as the joint distribution of these two objects. Hence, I obtain very direct information on the explanations of why star ratings might not have a significant impact on demand. On the

[19] In a different study ([67]), the authors also analyzed the effects of the SRS on the premiums set by insurers and the total welfare.

[20] Recently, there has also been a growing branch of literature that focuses on potential risk selection that can arise due to either the SRS or the QBP program on the demand or the supply side of the market ([65], [35], [34], [45]). My study does not focus on any type of selection.

supply side, I model firms to compete on more than one dimension, as in [44] and the more recent [73]. An innovative part of my study is the inclusion of a quality targeting error that prevents insurers from targeting exactly a certain quality level required to get a bonus. Lastly, I also allow firms to endogenize the information environment when they make their decisions. This allows me to investigate the effects of the financial incentives given to insurers separately from demand-side effects.

Further, my study integrates techniques from the behavioral ([33], [69]), marketing ([47], [70], [11]), and transport economics ([61], [23], [96], [24]) literature by collecting data via a survey design and using a conjoint analysis to elicit preferences. Using conjoint experiments to estimate preferences has the advantage of exogenously varying prices and star ratings resulting to demand estimates that are not contaminated by endogeneity issues that usually arise when estimating demand with standard revealed preference choice data.

With the exception of [52] in the health economics literature, this is the first paper that combines stated preference and revealed preference choice data to estimate a full equilibrium model in the Industrial Organization literature.[21] The novel part of my methodology is that combining the information and choice data from the survey with real data on MA plan choices and characteristics, I estimate two different random coefficient demand models - one

[21]Studies that combine stated and revealed preferences data are usually found in the behavioral, marketing, and transport economics literature. Some of these studies include [10], [19], [40], [93], [22], [78], [75].

on the survey data and one on the actual demand data - and I "map" these models into one another using a scale parameter. In this way, I estimate a real world demand model for Medicare Advantage plans that also incorporates the effects of imperfect information, while the aspects that are not identified from the real choice data are identified by the survey data.

Lastly, my study uses empirical tools from the learning Bayesian literature ([5], [26], [42], [27], [48]).

Imperial Insurance Company of Texas Traditional (HMO) (H2793-003-0)

Organization: Imperial Insurance Company of Texas, Inc

Estimated Annual Drug Costs: [?]	**Monthly Premium:** [?]	**Deductibles** [?] **and Drug Copay** [?] **/ Coinsurance:** [?]	**Health Benefits:** [?]	**Drug Coverage** [?] **, Drug Restrictions** [?] **and Other Programs:**	**Estimated Annual Health and Drug Costs:** [?]	**Overall Star Rating:** [?]	
Retail Annual: $1,032 **Mail Order** Annual: N/A	$0.00 Drug: $0.00 Health: $0.00 **Part B Premium Reduction: No**	Annual Drug Deductible: $0 Health Plan Deductible: $0 Drug Copay/ Coinsurance: $0 - $90, 33%	Doctor Choice: Plan Doctors for Most Services Out of Pocket Spending Limit: $4,000 In-network D V H	All Your Drugs on Formulary: **No** Drug Restrictions: **No** **MTM Program : Yes**	$3,270	Plan too new to be measured	Enroll

Aetna Medicare Plus Plan (PPO) (H5521-199-0)

Organization: Aetna Medicare

Estimated Annual Drug Costs: [?]	**Monthly Premium:** [?]	**Deductibles** [?] **and Drug Copay** [?] **/ Coinsurance:** [?]	**Health Benefits:** [?]	**Drug Coverage** [?] **, Drug Restrictions** [?] **and Other Programs:**	**Estimated Annual Health and Drug Costs:** [?]	**Overall Star Rating:** [?]	
Retail Annual: $1,560 **Mail Order** Annual: N/A	$0.00 Drug: $0.00 Health: $0.00 **Part B Premium Reduction: Yes**	Annual Drug Deductible: $345 Health Plan Deductible: $0 Drug Copay/ Coinsurance: $2 - $100, 26%	Doctor Choice: Any Doctor Out of Pocket Spending Limit: $10,000 In and Out-of-network $6,700 In-network D V H	All Your Drugs on Formulary: **No** Drug Restrictions: **No** **MTM Program : Yes**	$3,670	★★★★ 4 out of 5 stars	Enroll

Figure 1.1: Plan finder

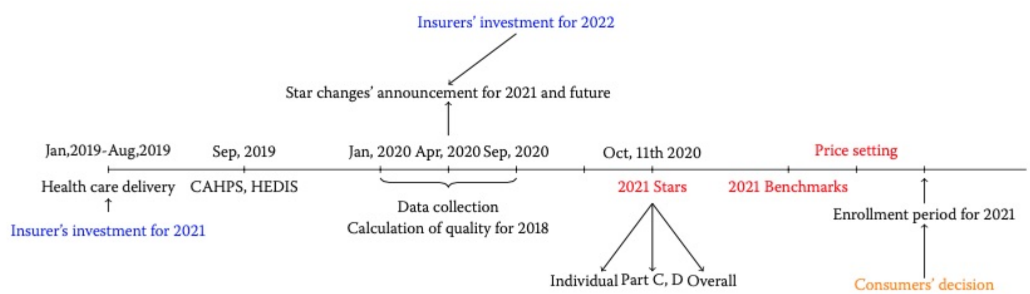


Figure 1.2: Example of a timeline between healthcare delivery, data collection, and star ratings release

Domain	Measure ID	Measure Name	Part C/D Weights	MA - PD Weights
Staying Healthy	C01	Colorectal Cancer Screening	1	1
	C02	Cardiovascular Care - Cholesterol Screening	1	1
	C03	Diabetes Care - Cholesterol Screening	1	1
	C04	Annual Flu Vaccine	1	1
	C05	Improving or Maintaining Physical Health	3	3
	C06	Improving or Maintaining Mental Health	3	3
	C07	Monitoring Physical Activity	1	1
	C08	Adult BMI Assessment	1	1
Managing Chronic Conditions	C09	Special Needs Plan Care Management	1	1
	C10	Care for Older Adults - Medication Review	1	1
	C11	Care for Older Adults - Functional Status Assessment	1	1
	C12	Care for Older Adults - Pain Assessment	1	1
	C13	Osteoporosis Management in Women who had a Fracture	1	1
	C14	Diabetes Care - Eye Exam	1	1
	C15	Diabetes Care - Kidney Disease Monitoring	1	1
	C16	Diabetes care - Blood Sugar Controlled	3	3
	C17	Diabetes care - Cholesterol Controlled	3	3
	C18	Controlling Blood Pressure	3	3
	C19	Rheumatoid Arthritis Management	1	1
	C20	Improving Bladder Control	1	1
	C21	Reducing the Risk of Falling	1	1
	C22	Plan All - Cause Readmissions	3	3
Member Experience with the Health Plan	C23	Getting Needed Care	1.5	1.5
	C24	Getting Appointments and Care Quickly	1.5	1.5
	C25	Customer Service	1.5	1.5
	C26	Rating of Health care Quality	1.5	1.5
	C27	Rating of Health Plan	1.5	1.5
	C28	Care Coordination	1.5	1.5
Member Complaints and Changes in the Health Plan's Performance	C29	Complaints about the Health Plan	1.5	1.5
	C30	Members Choosing to Leave the Plan	1.5	1.5
	C31	Health Plan Quality Improvement	5	5
Health Plan Customer Service	C32	Plan Makes Timely Decisions about Appeals	1.5	1.5
	C33	Reviewing Appeals Decisions	1.5	1.5
Drug Plan Customer Service	D01	Appeals Auto - Forward	1.5	1.5
	D02	Appeals Upheld	1.5	1.5
Member Complaints and Changes in the Health Plan's Performance	D03	Complaints about the Drug Plan	1.5	1.5
	D04	Members Choosing to Leave the Plan	1.5	1.5
	D05	Drug Plan Quality Improvement	5	5
Member Experience with the Drug Plan	D06	Rating of Drug Plan	1.5	1.5
	D07	Getting Needed Prescription Drugs	1.5	1.5
Drug Safety and Accuracy of Drug Pricing	D08	MPF Price Accuracy	1	1
	D09	High Risk Medication	3	3
	D10	Diabetes Treatment	3	3
	D11	Medication Adherence for Diabetes Medications	3	3
	D12	Medication Adherence for Hypertension	3	3
	D13	Medication Adherence for Cholesterol (Statins)	3	3

Table 1.1: Individual performance measures

	Star Ratings					
Year	≤ 2.5	3	3.5	4	4.5	5
2008 – 2011	0%	0%	0%	0%	0%	0%
2012 – 2013	0%	3%	3.5%	4%	4%	5%
2014	0%	3%	3.5%	5%	5%	5%
2015 – present	0%	0%	0%	5%	5%	5%

Table 1.2: Quality bonus payment percentages by star rating levels

	Star Ratings					
Year	≤ 2.5	3	3.5	4	4.5	5
2008 – 2011	75%	75%	75%	75%	75%	75%
2012	66.7%	66.7%	71.7%	71.7%	73.3%	73.3%
2013	58.3%	58.3%	68.3%	68.3%	71.7%	71.7%
2014 – present	50.0%	50.0%	65.0%	65.0%	70.0%	70.0%

Table 1.3: Rebate percentages

Chapter 2

Preliminary Analysis

In this chapter I introduce and describe the data I use to investigate the market and the policies of interest. The data come from publicly available data sources at the Centers for Medicare and Medicaid Services (CMS). Further, using standard discrete choice methods, I estimate demand for Medicare Advantage (MA) plans and provide evidence that star ratings do not impact beneficiary plan choices. Lastly, I provide preliminary evidence that the star ratings improvement observed in the market was mainly motivated by the financial incentives insurers received through the Quality Bonus Program (QBP).

2.1 Data

Every year CMS allows beneficiaries to compare characteristics of plans that are available in their location through a Plan Finder website. I use the Plan Finder database for the period 2008 – 2016 to compile information from various data sets on star ratings, premiums, benefits, and enrollment levels.

I extract data on overall star ratings for every existing contract per year. Figure 2.1 displays how the mean overall star ratings of all contracts has evolved over time. The market started experiencing an increase in star ratings in 2011 in anticipation of the QBP. The average overall star rating first exceeded 4 in 2015, the year when the QBP took its original and final form.

I also extract data sets on the premiums (Part B, C and D), the out-of-pocket costs beneficiaries may incur in the plans in which they enroll, indicators for drug, dental, vision, and hearing coverage, and indicators for Health Maintenance Organization, Preferred Provider Organization, and Private Fee For Service plan types for every plan and year. Moreover, I extract information on plan enrollment and Medicare eligibility levels by state and county and construct market shares for MA plans and Traditional Medicare (TM). From the constructed data set, I drop plans that are sponsored by employers and those that are designed for individuals who are "dually eligible" for Medicare and Medicaid. Plans in these categories operate under a different payment and benefit structure. I also drop plan-county observations with ten or fewer enrollees. CMS reports that these observations are created mostly by individuals who move outside of the plan's service area.

Table 2.1 presents summary statistics for each county-year-plan observation during the period 2008 – 2016. The total number of observations is 130607. There are 23743 county-years in total, each corresponding to a different market, 613 contracts, and 3890 plans across the nation. The average number of contracts per market is 1.56, and the average number of plans per market is 5.5. The average enrollment per plan is 378.3 with a high standard deviation, implying that there is wide geographic variation in the number of potential enrollees. The average star rating per contract is 3.7, and most contracts receive star ratings of 3.5 – 4.

Table 2.2 presents summary statistics on plan characteristics based on different levels of star ratings.[1] The data do not reveal any patterns between plan benefits and star ratings. On average, a beneficiary pays $36 on Part C premium, $31 on Part D premium, and $254 on out-of-pocket-cost for a 4-star plan. Higher-star plans are more expensive. This may contradict the fact that star ratings do not capture financial aspects of quality. However, this pattern can be attributed to the mechanics of the star ratings construction, as people who report to being satisfied in the Consumer Assessment of Healthcare Provider and Systems and the Healthcare Effectiveness Data Information Set surveys can be enrolled in higher-premium plans.[2]

CMS does not provide data on the weighted average summary rates that result in the discretized star ratings. Instead, it provides information on the algorithms it uses to generate them as well as data on the star ratings for each individual metric that composes the overall star rating of a contract.[3] I extract these data, and in combination with the algorithm information, I generate the weighted average summary rates. Figure 2.2 shows how the average summary rates I construct align with the star ratings for the year 2013. The horizontal axis represents the average summary rates as I construct them, and the vertical

[1]Information on plans receiving less than "2.5 star" ratings is not reported because of the limited availability of such plans in the data.

[2]Another explanation could be that insurers price star ratings.

[3]Section 1.1 of the appendix describes the main algorithm CMS uses.

axis represents the star ratings reported in the original data set. As expected, the shape of the plot resembles a step-wise function with jumps around the thresholds where a new star rating is assigned. The overlaps are due to data limitations that prevent perfect matching between the average summary rates and the star ratings data.

2.2 Investigating the demand side

In this section I investigate the demand for health insurance plans to quantify consumer willingness to pay for star ratings among other plan characteristics. I use a fairly standard nested logit specification for plan demand that has often been used in the literature ([91], [39], [29]) and I estimate demand for MA plans. I assume that consumers are perfectly informed with regard to the SRS and thus I treat star ratings as an additional plan characteristic.

I define a market as a county-year. Plans in every market are differentiated with respect to premiums, star ratings, and benefit design. In every market, consumer i makes a choice of a plan j, offered by insurer k, which maximizes her utility based on all the observed plan characteristics among all the available options. The nested logit specification will allow consumers to have correlated tastes within the same group of products. I divide the existing plans into two exhaustive and mutually exclusive groups, g: MA plans belong to group $g = 1$, and TM, which represents the outside option, belongs to group $g = 0$.

The utility beneficiary i receives from plan j in county m at time t is given by:

$$u_{ikjmt} = \delta_{kjmt} + \zeta_{ig}(\pi) + (1-\pi)\epsilon_{ikjmt}, \tag{2.1}$$

where

$$\delta_{kjmt} = \alpha p_{kjmt} + \beta r_{kt} + \gamma x_{kjmt} + \xi_k + \xi_{kjmt}. \tag{2.2}$$

In this specification, p_{kjmt} denotes the premium, r_{kt} the star ratings, x_{kjmt} the plan characteristics that refer to the benefit design (drug, dental, vision, and hearing coverage), and ξ_{kjmt} the unobserved to the econometrician plan characteristics that usually refer to provider networks and plan advertising. Finally, ϵ_{ikjmt} is a zero-mean i.i.d. stochastic error term following the Type I extreme-value distribution across plans and consumers, and ζ_{ig} is another error term drawn from a distribution with parameter π such that $\zeta_{ig}(\pi)+(1-\pi)\epsilon_{ikjmt}$ follows a generalized extreme value distribution.

As shown in [12], this will yield the market share, s_{kjmt}, of a plan

$$s_{kjmt} = \frac{e^{\delta_{kjmt}}}{1+\sum e^{\delta_{kj'mt}}}, \tag{2.3}$$

where the utility a beneficiary receives from the outside good u_{i0mt} has been normalized to 0. Therefore, the final equation for estimation is

$$ln(s_{kjmt}) - ln(s_{0m}) = \delta_{kjmt} + \pi ln(s_{g|i}). \tag{2.4}$$

Table 2.3 presents the estimates. Specifications (1) and (2) present OLS estimates with and without time fixed effects, respectively. Specifications (3) and (4) present IV estimates with and without time fixed effects, respectively. All specifications include contract fixed effects to account for the fact that there might be contract specific effects that are not captured by the characteristics the current specifications control for. Consistent with OLS estimates on the price being biased towards zero, the IV price coefficients are larger in magnitude than the OLS coefficients. For this reason, I focus on the last two specifications. All specifications of interest give a relatively low coefficient on star ratings compared to other plan characteristics. The value consumers assign to an extra star per month varies between \$3 – \$11.

Table 2.4 also presents specifications coming from a standard random coefficient model that is estimated following the standard [14] algorithm. The first three specifications present estimates without including instrumental variables. For the same reason as previously, I focus on the last specification that includes instrumental variables. The estimates convey the same message. Consumer willingness to pay for an extra star rating amounts to \$1.

All demand estimates suggest that consumers tend to value drug coverage significantly more than every other plan characteristic. This result is not new. It is fairly accepted in the health insurance literature that consumers have high willingness to pay for drug coverage. However, it is worth to notice that consumer willingness to pay for star ratings is close to, and in most cases lower than, their average willingness to pay for plan characteristics such as dental, vision, or hearing coverage. This observation arises concerns about the reason why this might happen. Is it because the star ratings information does not get to consumers? Or, is it because star ratings are not very important to consumers compared to other plan characteristics?

2.3 Investigating the supply side

In this section I investigate the supply side of the market analyzing how the QBP affected the star ratings evolution.

The first observation was made at the beginning of the chapter based on figure 2.1. The figure shows that the increase in the star ratings started one year before the QBP was implemented. This observation suggests that the QBP contributed to the star ratings improvement significantly more than the information provision through the SRS did in the previous years.

Figure 2.3 displays the density distribution of the continuous quality before and after the QBP. The graph indicates that after the QBP was implemented the modes of the distribution are closer to the levels of quality where the reimbursements would occur. Interestingly, we do not see bunching exactly around the critical thresholds that should be achieved for different star rating levels. This observation implies that even though insurers might target specific quality levels, they cannot do so perfectly. If insurers were able to target the quality levels they would like to achieve perfectly, we would see clear bunching around the corresponding critical thresholds.

Investigating the evolution of the star ratings in more depth, figure 2.4 displays how the weighted-by-enrollment distribution of the average summary rates has evolved over time.[4] The distributions are multi-modal, with the modes around the thresholds where a new star rating is assigned. Interestingly, after the implementation of the QBP in 2012 the highest modes of the distributions are realized close to the thresholds where the plan reimbursements occur. Those observations imply that insurers tend to target specific ratings and that they are effectively motivated by the financial incentives.[5] Noticeably, there are insurers who receive more than "4 star" ratings, which might seem nonsensical since the bonuses are not higher for 5-star plans. This observation reflects the fact that 5-star plans qualify for extra benefits compared to lower-star plans; for example, they can enroll beneficiaries at any time during the year.

Figure 2.5 displays kernel density estimates around the critical thresholds that lead to the corresponding star rating levels for all years of data together. A common pattern observed in most cases is a multi-modal distribution with modes to the left and to the right of the critical thresholds. This observation implies that knowing the star assignment algorithm used by CMS, insurers try to ensure they will get to the necessary thresholds to achieve a specific star rating level, but they cannot target perfectly. Interestingly, the masses drop around the 3.75, 4.25 and 4.75 critical thresholds. If insurers did not target specific levels of quality, we would expect smoother distributions

[4] I exclude plans that have entered and/or exited the market during the period of interest to avoid concerns that entry/exit of plans affect the observed patterns.

[5] During 2012 − 2014, the modes of the distributions are almost equally spread around the different thresholds, whereas after 2014, when the QBP program took its final form, higher masses of the distributions are realized around the 4-star rating threshold.

around the thresholds. Conversely, if insurers did target specific thresholds, we would expect excess masses around the thresholds. It is interesting to notice that uneven patterns are observed mostly around the thresholds leaning to 3.5 and 4 star ratings suggesting that insurers try to target the levels that will secure them the financial bonuses.

Further, I conduct [68] "manipulation" tests–widely used in the Regression Discontinuity econometric literature–around critical thresholds. It is important to mention that "manipulation" in this case does not imply "gaming". Instead, it simply describes a situation where insurers endogenously make decisions on the quality levels they desire to achieve. The null hypothesis of these tests is the one of "no manipulation". All the tests were conducted using the "rddensity" stata command. Figure 2.6 displays the results. Most point estimates lie outside the confidence intervals suggesting that the tests failed the null hypothesis of no manipulation.[6]

In addition, figures 2.7, 2.8, 2.9, 2.10, 2.11, 2.12 display results of [68] tests around different thresholds for periods before and after the QBP was implemented. Again the tests failed the null hypothesis of no manipulation in almost all cases. It is interesting, however, to notice that around the critical thresholds where insurers would be reimbursed with bonuses, the tests do not fail the null hypothesis of no manipulation before the QBP was implemented, but they fail it after.

[6] A test around the 2.25 critical threshold was not doable due to lack of sufficient observations around the threshold.

2.4 Conclusions

Overall, the analysis of this chapter suggests that consumer willingness to pay for star ratings is relatively low compared to other plan characteristics. This observation can partially explain why the market does not experience star ratings improvement during the first years that the SRS was implemented. If consumers are not willing to pay for higher star ratings, then insurers do not have any incentive to invest in them. In addition, the shift of the density distribution of the continuous quality over time combined with the "no manipulation" tests results suggest that insurers were incentivized by the financial incentives the QBP provided.

Before making definitive conclusions a few caveats should be taken into account. First, low levels of demand responses to star ratings do not necessarily mean that consumers do not value star ratings. Instead, this observation arises the question of whether consumers are aware of the SRS and the information it provides. In the available data consumers who are not aware of the SRS are observationally equal to consumers who do not value star ratings. Importantly, consumers who are not aware of the SRS might value the star ratings had they been aware of it. Even more interestingly, they might have different levels of willingness to pay for star ratings from consumers who are currently aware of the SRS. Assuming perfect information can lead to demand estimates that represent mixed values of two different consumer types. But, in a study that analyzes the welfare effects of an information provision program, it is important to distinguish different types of consumers with respect to their awareness levels and their willingness to pay so that the right policy implications to be made.

Second, the current demand model treats star ratings as an additional plan characteristic. However, star ratings represent (i.e. they are a signal of) a specific dimension of quality that refers to customer satisfaction and clinical outcomes. Concluding that consumers do not value star ratings can lead to misleading interpretations in regards to whether consumers value star ratings as signals of a specific dimension of quality or whether consumers value this specific dimension of quality per se. Distinguishing between the two is important for the right welfare analysis and policy recommendations.

Third, although the shift of the continuous quality distribution suggests quality improvement due to the QBP, it is not clear whether this shift is completely due to the financial incentives. Since the distributions are weighted by enrollment, one could conclude that there is demand for higher-star plans and thus demand side effects too. However, the high levels of inattention ([77], [54]) that plague this market do not allow for a clear conclusion. And if the high enrollment levels of the high-star plans do not reflect preferences, a question that refers to the justification of the cost the market incurs to attain significant quality improvement arises. Since distinguishing both the demand and supply side effects of the two policies is important from a welfare perspective and since clear conclusions cannot be made from the above analysis, building a structural model is recommended.

The analysis of the next chapters shed light to the questions/concerns stated above. In the next chapter, I describe and analyze the results coming from a survey I conduct. The goal of the survey is to identify the percent of the MA beneficiaries that are not aware of the SRS and their corresponding preferences given their type. After I conclude my survey analysis, I build and estimate a full structural demand and supply equilibrium model that analyzes the relative impacts of both policies, SRS and QBP, on welfare. Importantly, on the demand side, I treat star ratings as signals of the quality dimension that refers to clinical outcomes and customer satisfaction. In addition, on the supply side, I model insurers to compete in two dimensions, price and star ratings, incorporating a part of uncertainty in their decisions that does not let them choose star ratings perfectly.

	Mean	SD	25 %	Median	75 %
Key plan variables					
Monthly Part B premium ($)	100.501	8.622	96.4	101.007	104.9
Monthly Part C premium ($)	28.818	36.053	0	15.2	46.5
Monthly Part D premium ($)	22.861	20.826	0	22	34.8
Other plan characteristics					
Overall star rating	3.705	0.609	3.5	3.5	4
Drug	0.760	0.426	1	1	1
Dental	0.330	0.470	0	0	1
Vision	0.813	0.389	1	1	1
Hear	0.667	0.471	0	1	1
Plan types					
hmo	0.295	0.456	0	0	1
ppo	0.277	0.447	0	0	1
pffs	0.426	0.494	0	0	1
Plan-level enrollment	378.309	1477.609	29	72	231
Market-level MA share	0.197	0.127	0.097	0.172	0.269
Total # of markets	23743				
Total # of contracts	613				
Total # of plans	3890				
# of contracts/ market	1.560	0.859	1	1	2
# of plans/ market	5.500	4.819	2	4	7
# of plans/ contract	1.642	0.906	1	1	2
Monthly payment ($)	800.315	141.901	699.435	781.361	884.233
N = 130607					

Table 2.1: Summary statistics

	Star Ratings					
Variable	2.5	3	3.5	4	4.5	5
Part C premium ($/month)	16.094	18.832	32.794	36.442	44.981	68.849
Part D premium ($/month)	14.252	17.603	27.440	31.161	32.021	34.345
OOPC ($/month)	221.807	211.391	233.032	253.633	228.673	220.029
Supplemental Coverage						
Prescription drugs	.757	.818	.860	.824	.839	.683
Dental	.412	.271	.284	.351	.482	.106
Vision	.911	.814	.845	.875	.884	.953
Hearing	.761	.607	.492	.432	.626	.858
Plan types						
HMO	.280	.254	.312	.305	.522	1.000
PPO	.165	.365	.483	.438	.477	0.000
PFFS	.553	.379	.203	.256	.000	.000
Market shares	.036	.048	.054	.052	.061	.048
Observations	3486	11345	17663	14800	12724	1104

Table 2.2: Plan characteristics by star ratings

Variable	(1)	(2)	(3)	(4)
α	-.001***	-.001***	-.023***	-.026***
	(.000)	(.000)	(.001)	(.001)
β	-.023***	.104***	.062***	.302***
	(.009)	(.009)	(.018)	(.016)
γ_{drug}	.201***	.183***	1.292***	1.451***
	(.008)	(.008)	(.046)	(.041)
γ_{vision}	.069***	.084***	.209***	.238***
	(.010)	(.010)	(.019)	(.017)
γ_{dental}	.073***	.090***	.549***	.670***
	(.007)	(.007)	(.024)	(.021)
γ_{hear}	.010	-.021*	.605***	.660***
	(.010)	(.012)	(.033)	(.029)
ln(plan MA share)	.817***	.826***	.684***	.676***
	(.002)	(.002)	(.007)	(.006)
Willingness to pay (in$/month) for:				
an extra star	-23	104	3	11
drug covg	201	183	56	56
vision covg	69	84	9	9
dental covg	73	90	24	25
hear covg	10	-21	26	25
Mean of dependent variable	-1.882***	-2.104***	-2.894***	-3.546***
	(.036)	(.035)	(.071)	(.083)
Mean own-price elasticity	-.369	-.318	-4.344	-4.838
Year FEs	Y	N	Y	N
Contract FEs	Y	Y	Y	Y
IVs	N	N	Y	Y
Observations	61146	61146	61146	61146
Adj. R^2	.753	.743	.752	.743

Table 2.3: Demand estimates - Nested Logit

Variable	(1)	(2)	(3)	(4)
α	-.005	-.007	-.008	-.055
β	.146	.279	-.050	.060
σ_α	.0007	.007	0.006	.044
σ_β	.002	.016	.002	.011
γ_{drug}	1.268	1.144	1.318	1.877
γ_{vision}	.108	-.610	.0715	.215
γ_{dental}	.140	.007	.119	.420
γ_{hear}	-.014	-.048	.070	.270
Willingness to pay (in$/ month) for:				
an extra star	29.2	39.8	-6.2	1.0
drug covg	253.6	163.4	164.8	34.1
vision covg	21.6	-76.2	8.8	3.9
dental covg	28.0	1.0	14.8	7.6
hear covg	-2.8	-6.8	8.7	4.9
Mean of dependent variable	-4.605	- 19.729	-3.113	21.625
Mean own-price elasticity	-.663	-1.276	-1.131	-9.578
Year FEs	N	Y	Y	Y
Contract FEs	N	N	Y	Y
IVs	N	N	N	Y
Observations	61146	61146	61146	61146

Table 2.4: Demand estimates - Random Coefficients

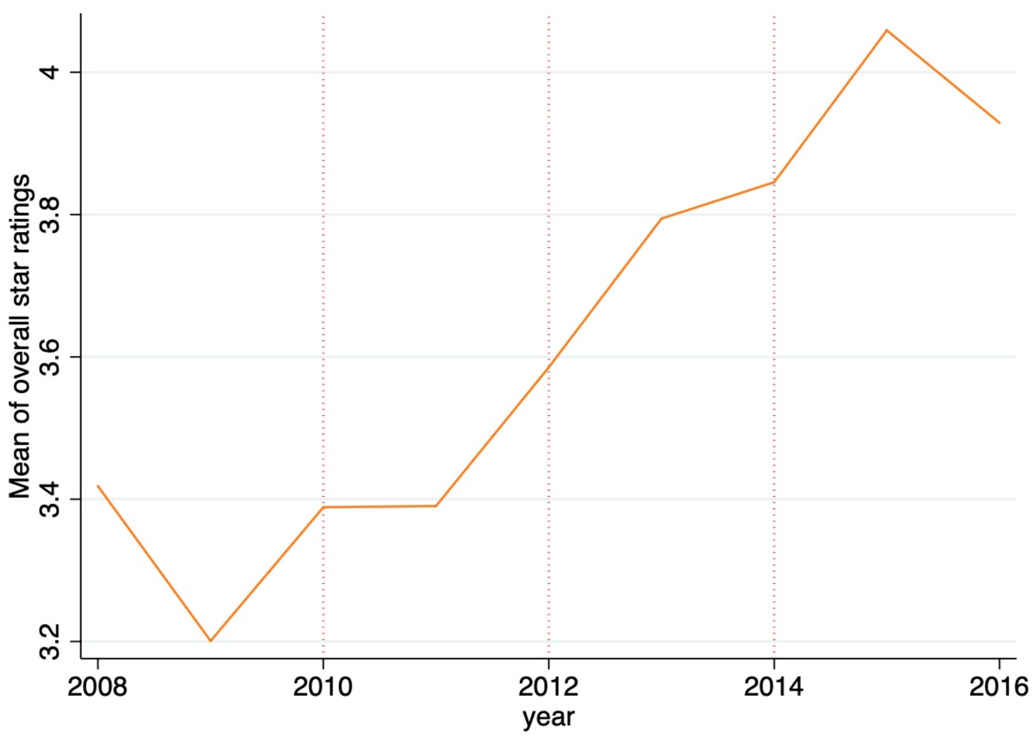


Figure 2.1: Average star ratings evolution over time

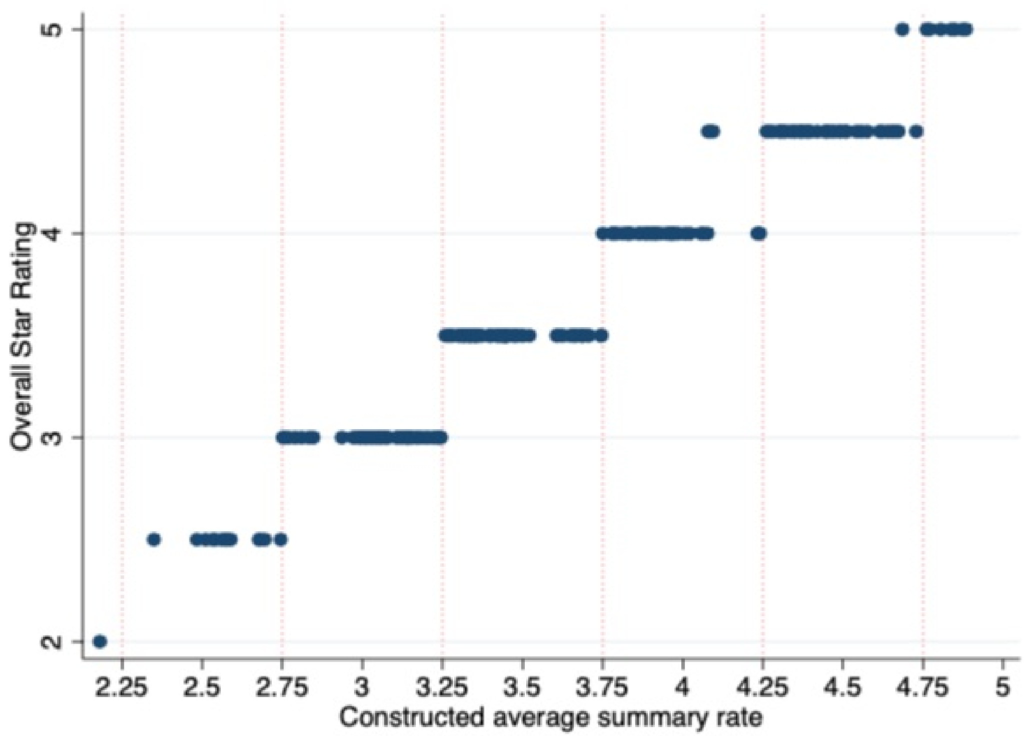


Figure 2.2: Constructed average summary rates

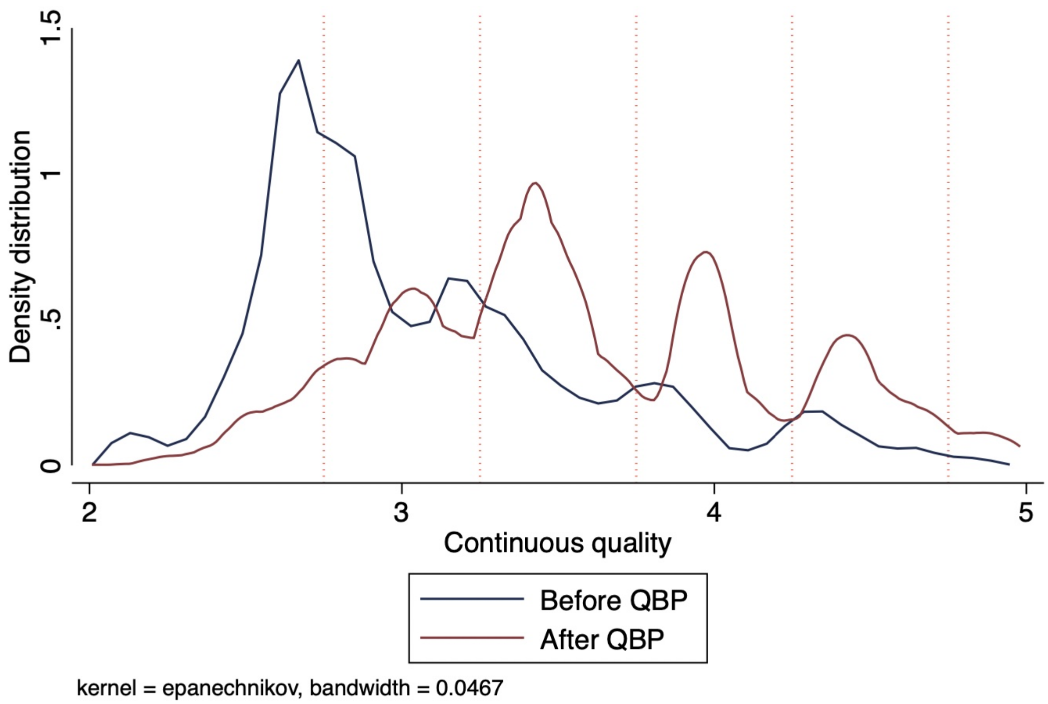


Figure 2.3: Density distribution of continuous quality before and after QBP

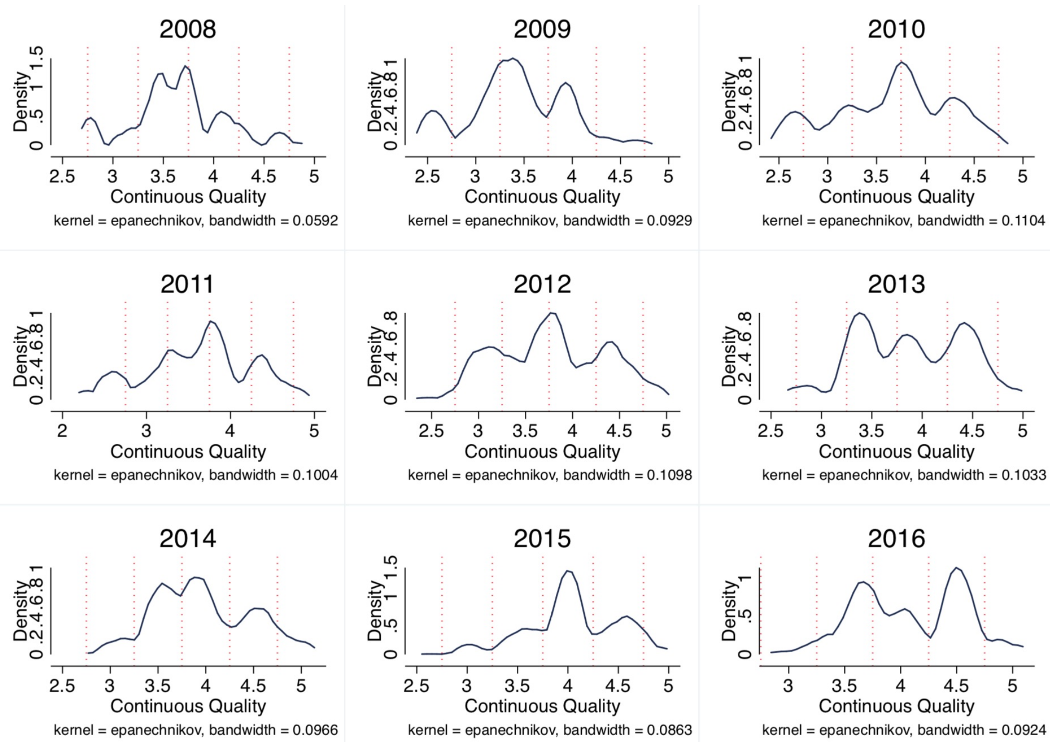
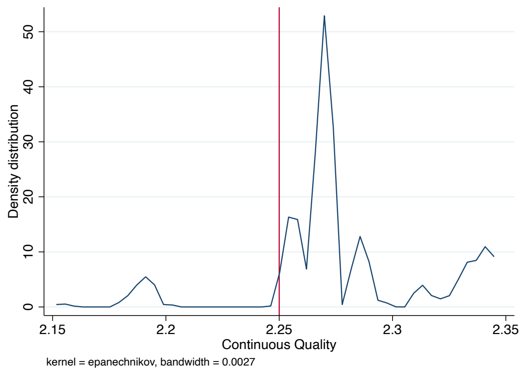
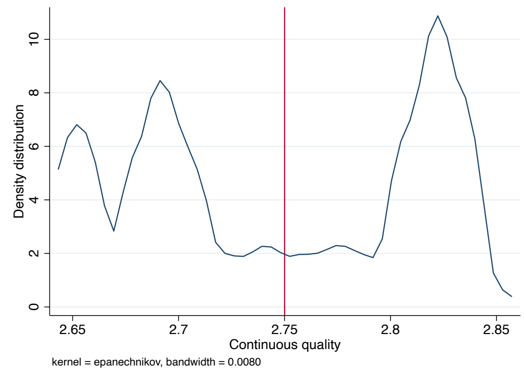


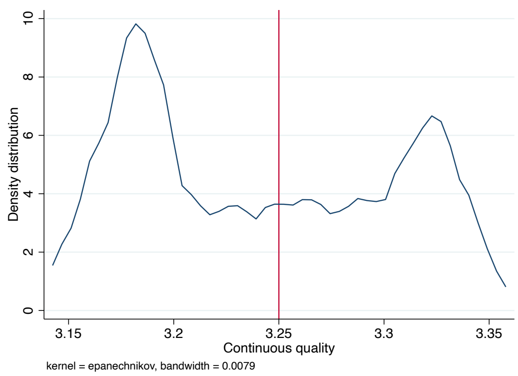
Figure 2.4: Distribution of continuous quality over time



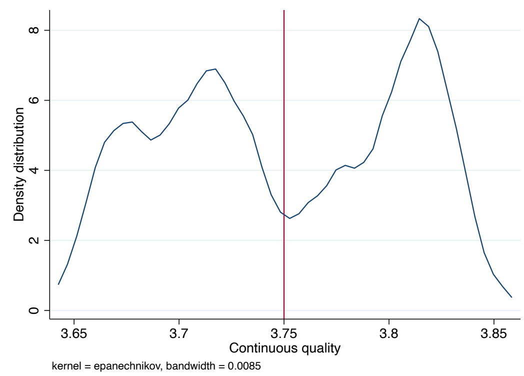
(a) Overall star = 2.5



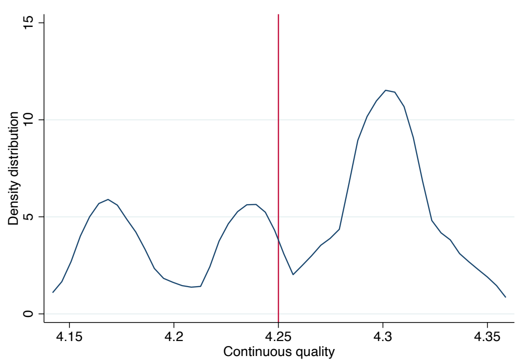
(b) Overall star = 3



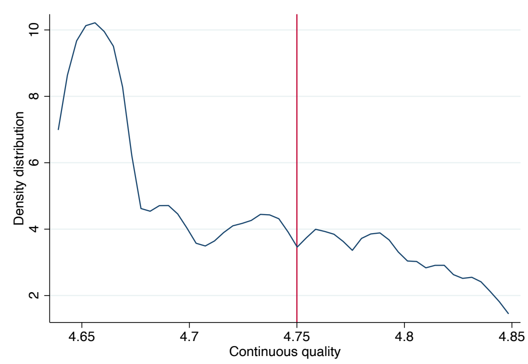
(c) Overall star = 3.5



(d) Overall star = 4

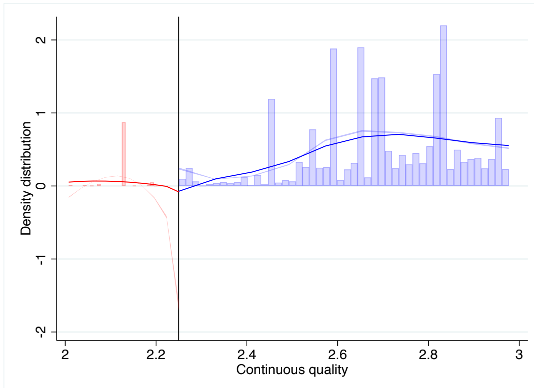


(e) Overall star = 4.5

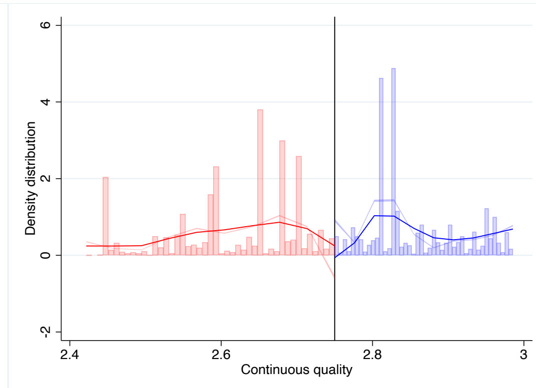


(f) Overall star = 5

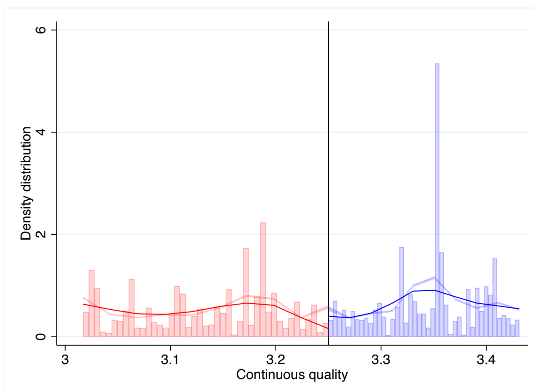
Figure 2.5: Kernel density estimates around different critical thresholds



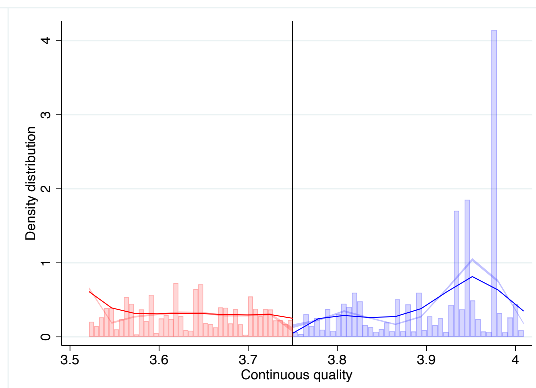
(a) Overall star = 2.5



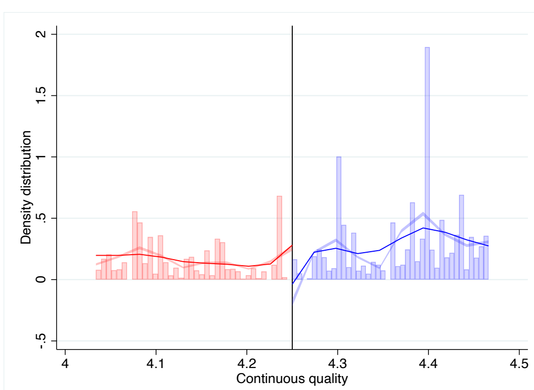
(b) Overall star = 3



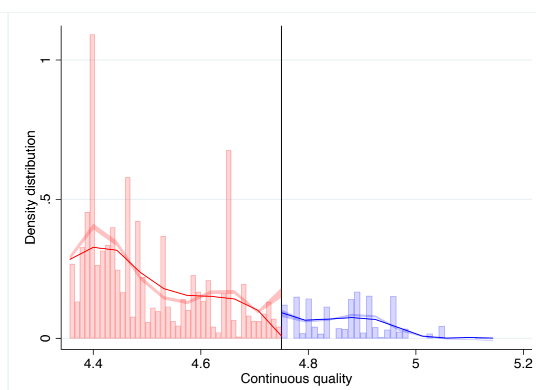
(c) Overall star = 3.5



(d) Overall star = 4

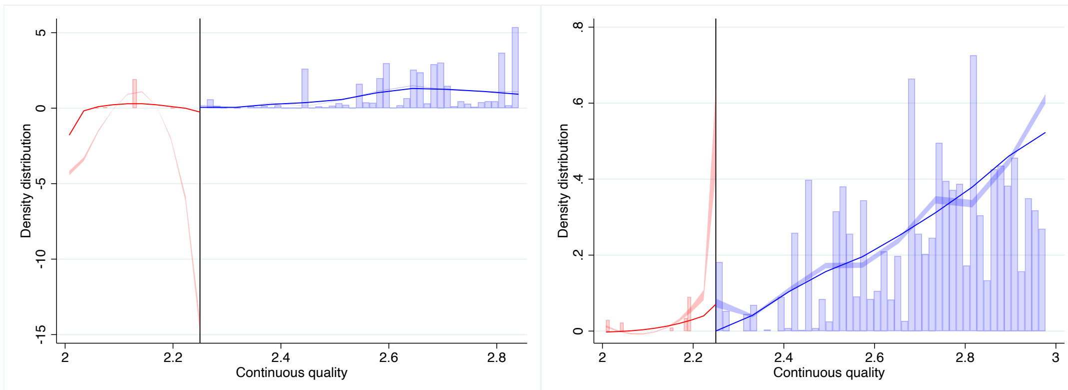


(e) Overall star = 4.5



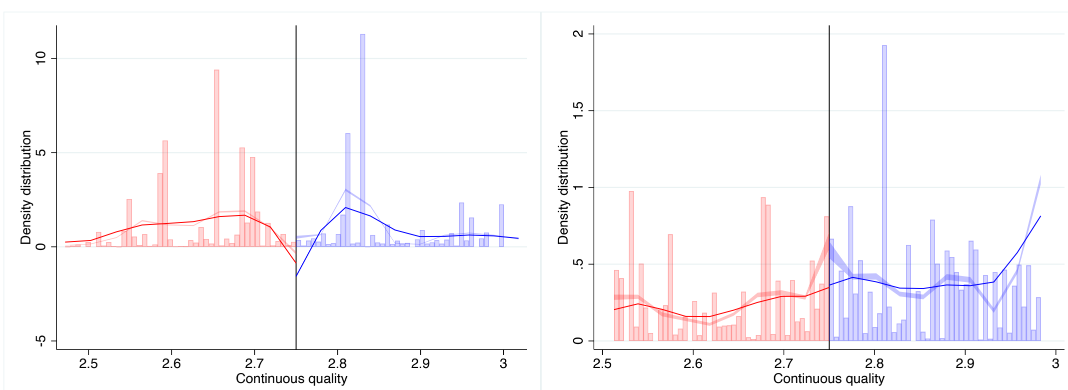
(f) Overall star = 5

Figure 2.6: [68] McCrary manipulation test estimates



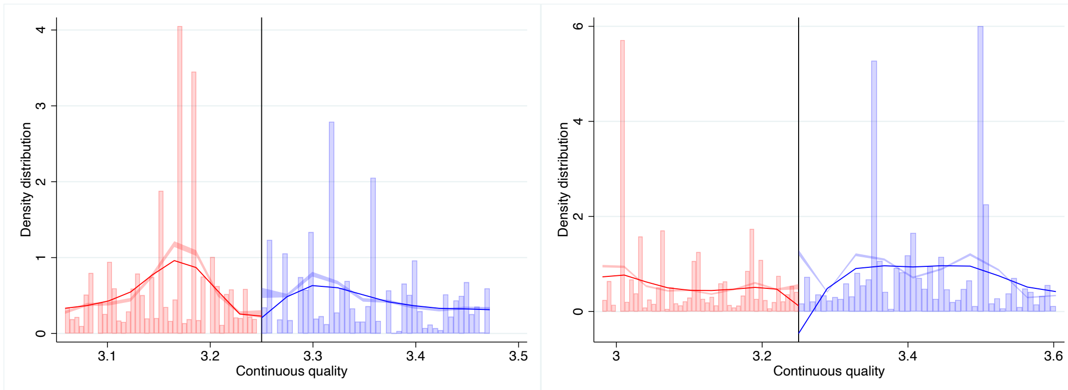
(a) Before QBP (b) After QBP

Figure 2.7: [68] McCrary manipulation test estimates at critical threshold 2.25



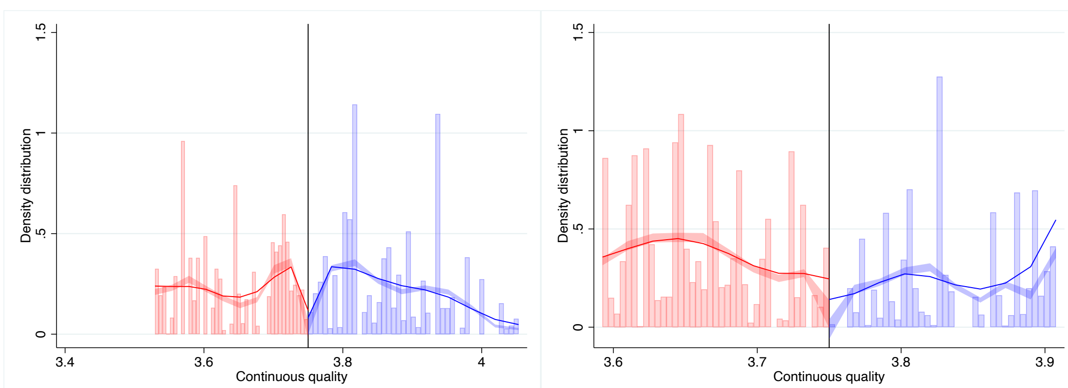
(a) Before QBP (b) After QBP

Figure 2.8: [68] McCrary manipulation test estimates at critical threshold 2.75



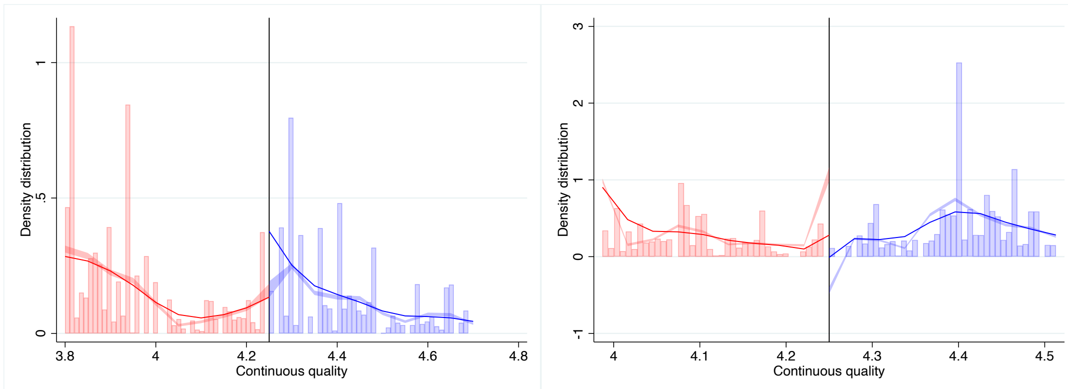
(a) Before QBP (b) After QBP

Figure 2.9: [68] McCrary manipulation test estimates at critical threshold 3.25



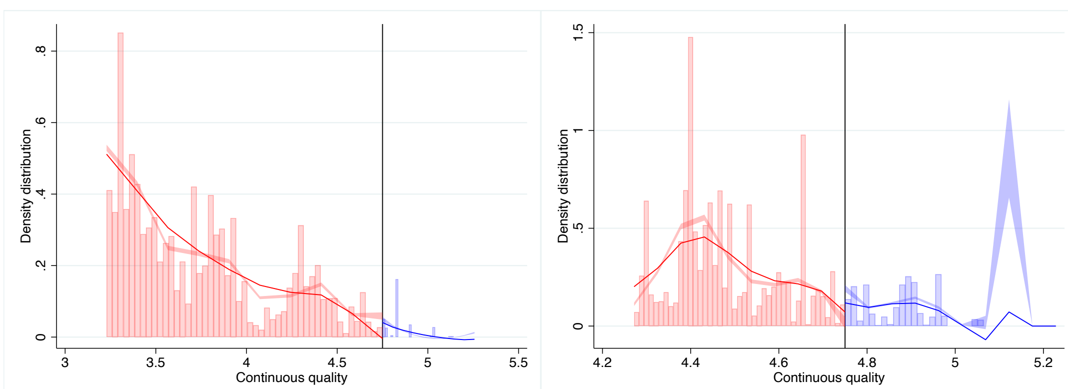
(a) Before QBP (b) After QBP

Figure 2.10: [68] McCrary manipulation test estimates at critical threshold 3.75



(a) Before QBP (b) After QBP

Figure 2.11: [68] McCrary manipulation test estimates at critical threshold 4.25



(a) Before QBP (b) After QBP

Figure 2.12: [68] McCrary manipulation test estimates at critical threshold 4.75

Chapter 3

Survey

In this chapter, I describe and analyze an electronic survey I designed and ran targeting the Medicare Advantage (MA) population. The survey yields unique information on consumer awareness with respect to the Star Rating System (SRS). Within the survey, I also conduct a conjoint analysis to elicit consumer preferences for star ratings. In this way I recover consumers' willingness to pay for star ratings given their *awareness* status, and also obtain clear explanations of the reasons that prevent star ratings from affecting demand for MA plans. Lastly, through an additional set of questions, I investigate what are the most important factors that affect consumer plan choice decisions.

3.1 New Survey Data on Medicare Population

The survey was executed by Qualtrics, which is a privately held experience management company that provides a survey platform/software for collecting and analyzing data for market research, customer satisfaction, loyalty, etc. Qualtrics has 6m members in N. America. It recruits respondents by providing them an incentive, the strength of which depends on the survey length and the target-level acquisition difficulty.

As the population of interest was MA beneficiaries, I stratified the sample by oversampling the MA population. The survey was distributed to 624 nationally representative respondents (456 MA and 168 Traditional Medicare (TM) beneficiaries) in two roll-outs; one in August 2019 and another in January 2020. The purpose of the two roll-outs was twofold. First, the initial roll-out served the purpose of a soft launch that could reveal potential problems; no substantial problems arose, and therefore no major changes were made in the second roll-out. Second, rolling out the survey in two different periods allowed me to compare the SRS awareness levels in periods closer and further away from the annual enrollment period.[1]

Respondents gained access to the survey by invitation only. The survey consisted of five sets of questions.[2] In the first set of questions recruits went through a layer of validation (mainly MA age eligibility) to ensure that they represented the target population. In the second set of questions I investigated consumer awareness with respect to the SRS. In the third set of questions I conducted the conjoint experiments to estimate the respondents' preferences with respect to star ratings. In the fourth set of questions I investigated the general process consumers follow before they make their insurance plan choice. Finally, in the fifth set of questions I collected demographic information of the

[1] In the following sections I show that there were not major differences in the results between the two periods.

[2] For the full questionnaire of the survey see section 2.2 of the appendix.

respondents.

After the data collection, I conducted a series of tests to ensure that there would not be gibberish responses in the final data. First, I excluded respondents whose IP address was not in the U.S. Second, I excluded respondents who did not spend sufficient time on the survey.[3] Finally, I excluded respondents who gave "meaningless" answers to "open-ended" questions.[4] Overall, I ended up with a total of 618 respondents. Out of this sample, none of the respondents dropped out during the survey.

Table 3.1 shows the length of the total duration of the survey in minutes as well as the time (in seconds) respondents spent on specific questions. Respondents spent 9.18 minutes on average to answer the entire survey. The main filtering question took them on average 36 seconds, while the question referring to their SRS awareness took them 17.5 seconds. They spent an average of 19 seconds on the first question of the conjoint experiment and an average of 7.5 seconds on the last one. This implies that, after the first conjoint experiment question, respondents became familiar with the concept of the experiment. Lastly, they spent almost a minute on the set of demographics questions. Table 3.2 also shows the corresponding durations split by the two roll-outs of the survey. The main difference in the results is the average

[3]Specifically, I excluded respondents who spent less than a percentile of the duration of the full survey or of the question referring to the coverage they had (TM vs MA).

[4]"Meaningless" responses were those that did not make sense or were unintelligible. Responses that were expressing any kind of emotions about Medicare or the survey itself were not excluded from the sample.

time respondents spent to answer the survey. In the second round of the survey respondents spent on average one min lower compared to the first. This difference reflects an additional set of questions included in the first round of the survey that required respondents more time to think. In the second round I excluded this set, because it did not provide meaningful results.

3.1.1 Filtering respondents

In this section, I describe the first set of questions respondents had to answer to ensure that the sample represented the target population. I constructed a set of criterion-questions that respondents had to answer before moving to the main part of the survey. Respondents who did not satisfy the necessary criteria were dropped from the sample.

The first eligibility criterion was age. Respondents below the age of 65 were excluded from the sample. Respondents were also asked to report the state of their current residence to ensure that the sample was nationally representative. Figure 3.1 shows the geographic variation of the survey population. The figure suggests that the sample is nationally representative. States in which there were not responses were Delaware and Wyoming. Also, respondents that reported they lived in Alaska were excluded from the sample, as this group of the population is also excluded from the sample of my main analysis.

Further, respondents had to report the main source–TM or MA–of their health insurance coverage. Respondents were provided example insurance cards explaining the main differences between TM and MA and they had to choose the option that applied to them. A third option, "Other", was also included to account for the remaining cases. Figure 3.2 shows the exact question they received. Respondents who chose "Other" were excluded from the sample. Further, respondents who reported they received coverage from their current or previous employers and/or from Veterans' programs were excluded from the survey, as these beneficiaries face different sets of plan options and benefits. Disabled population should also be excluded from the sample. However, the addition of a question investigating "disability" status would complicate the screening process due to the institutional complications of the industry and the properties that qualify a beneficiary to be classified as "disabled".[5] Overall, almost 50% of the respondents who satisfied the criteria of age and main eligibility did not satisfy the rest of the criteria (mainly due to their "employer-sponsored" insurance coverage) and were dropped. After respondents were filtered, they received access to the main part of the survey, where they had to answer a series of questions related to their SRS awareness and their preferences for star ratings. A set of a few extra questions, necessary for robustness analysis, followed.

3.1.2 Collecting demographic information

Collection of demographic information was necessary to ensure that the sample represented well the Medicare population. I asked respondents

[5] An additional reason to exclude such a question was also the possibility that it could affect respondents' answers and would question the validity of the survey.

to report their gender, ethnicity, marital status, monthly spending, education level and health status providing them various options. The options I provided were similar to the options provided by the Medicare Current Beneficiary Survey (MCBS), which is a rotating panel survey that tracks representative Medicare beneficiaries on an annual basis.[6] Lastly, I asked respondents the zip-code of their current residence.

Table 3.3 compares demographic information between my survey's sample population and MCBS's sample population, both split by MA and TM enrollment. The population is slightly younger in my sample than in the MCBS sample. My sample does not have considerable Black and/or Hispanic representation in either the MA or the TM populations compared to the MCBS sample. The education attainment is similar between the two populations. In my survey, I did not include the option "college attainment" in the "education-related" question, and thus the "college attainment" variable is subsumed in the "bachelor degree" variable. Respondents in my survey reported that they are slightly healthier compared to what MCBS respondents report. Lastly, "spending" refers to the beneficiaries' annual spending on housing, utility bills, food, transportation, healthcare, and other common leisure activities. On average, "spending" is lower for the MA population. Such a variable does not exist in the MCBS data. Instead, there is information on the annual income

[6]MCBS provides information on beneficiaries' annual enrollment decisions and their relevant demographic information. Although I do not use this data set in my analysis, I use this information to show the main differences between my survey population and the bigger MCBS population.

of Medicare beneficiaries. Overall, such differences are not surprising since the techniques and methodologies followed for the data collection are different between MCBS and Qualtrics.

Table 3.4 presents demographic information of the survey sample as it arose in the two roll-outs separately. Overall, the samples seem to be similar between the two roll-outs. The main difference is with respect to age. Beneficiaries in the first roll-out are on average 71-72 years old, while beneficiaries in the second roll-out are on average 67 years old. This difference is due to the fact that in the second round of the survey, Qualtrics was able to target MA population that was specifically between 65-70 years old.[7] Concerns about the sample not being representative should not arise, since if younger beneficiaries are not aware of the SRS, we would not expect older beneficiaries to be more aware.

3.1.3 Eliciting information on Information

To elicit information on the level of consumer *awareness* with respect to the SRS, after I briefly described to respondents the SRS and its scope I directly asked them whether they knew anything about it. To ensure that the information I provided about the SRS on this question would not affect respondents' answers, I followed the exact description of the SRS PlanFinder website gives actual consumers. Figure 3.3 displays the exact question as it

[7]Beneficiaries in this age group had a higher chance to be first time enrollees. As first time enrollees are considered being "active" choosers, focusing on them can provide better demand estimates.

appeared in the survey.

Table 3.5 presents the results. Overall, 77.4% of the entire Medicare population reported they were *unaware* of the SRS.[8] Comparing the awareness levels between the TM and MA populations, 17.3% and 32.1% reported that they were *aware* of the SRS, respectively. This difference reflects that MA beneficiaries tend to be more informed about the MA market. Overall, the results indicate that, regardless of the efforts and the sources the government allocates in the collection of the necessary information as well as the construction and the release of the star ratings, consumers are not aware of them.

Table 3.6 also shows how the results differ between the two roll-outs of the survey. It turns out that consumers are more aware in the second roll-out than in the first one, which reflects the fact that immediately after the enrollment period had ended consumers were more informed. An additional factor that could have affected this outcome is the younger sample of the second roll-out. However, the results do not differ much in magnitude, and the patterns of *awareness* are very similar between the roll-outs, with most of the population reporting to be *unaware* of the SRS.

Table 3.7 displays how *awareness* varies between the TM and MA pop-

[8]This result almost coincides with the result from another survey conducted by HealthMine in August 2018, indicating that 78% of the Medicare population was unfamiliar with the SRS. For their results, see `https://www.prnewswire.com/news-releases/only-22-of-medicare-advantage-members-are-familiar-with-star-ratings-of-those-stars-helped-half-choose-a-plan-healthmine-survey-300690119.html`.

ulations given their demographics. Both types–*aware* and *unaware*–seem to be similar. Respondents who reported they were *unaware* of the SRS were slightly older, slightly less educated, and less healthy. Also more females reported to be *unaware*. The data set does not provide enough variation to compare the populations with respect to race and ethnicity. Lastly, tables 3.8 and 3.9 display the same information between populations in the two roll-outs separately. No major differences are observed between the groups.

3.1.4 Information on Preferences–Conjoint Analysis

To elicit information on preferences for star ratings, r_{kmt}, after I separately identified the *aware* from the *unaware* consumer types I conducted a conjoint analysis following [11]. Conjoint analyses have been widely used in marketing research to measure consumer preferences and to forecast demand for components of a prospective product or service. The idea behind a conjoint analysis is that respondents are invited to make a series of choices requiring a trade-off between hypothetical product options.

The experiment was executed as follows. First, I provided all consumer types–ex ante *aware* and *unaware*–with a message to remind them what the SRS and its scope were and to give them information on the distribution of the star ratings. In this way, I ensured that everyone understood the signaling role the star ratings played and also realized the average level of star rating a plan gets. Figure 3.4 displays the exact message respondents received. Second, assuming that consumers prefer lower prices and higher star ratings, I generated a number of menus (choice sets) that randomly presented a trade-off between prices and star ratings. In reality, consumers face a rich set of options with each option characterized by many plan characteristics. However, to keep the experiment tractable and the length of the survey short, I gave respondents a simplified version of choice sets with only two options and two characteristics–prices and star ratings–for each option. Each respondent was provided with a set of 4 or 8 different menus, having to choose their preferred option for each menu. Figure 3.5 shows an example question for a menu respondents were presented. Table 3.10 shows all the possible values of the attributes presented in the experiments.

Figure 3.6 further displays the price and star differences that were presented in each menu provided to survey respondents. More specifically, a point in the graph represents the difference of the star ratings between the options that were available in a menu along with the corresponding difference of the prices between those two options.[9] Lastly, figures 3.7 and 3.8 display the price and star ratings variation in the menus that were provided to the respondents, respectively. All figures provide evidence that there was significant price and star rating variation in the menus of the conjoint experiments.

A concern that might arise regarding the experiment is the trade-off between the realism of the options and the ability of respondents to report their preferences in a short time. If the characteristics provided are too numerous

[9]For example, point (1,40) represents a menu where the difference in the star ratings of the two available options was 1 and the price difference between these two options was $40.

(generally more than 6 in a given choice) respondents tend to choose the ones that look simpler and more important to them without paying attention to the rest. In my conjoint analysis there are not such concerns, since I included only two attributes for each choice I provided to respondents.

3.1.4.1 Estimating preferences

A natural approach used in conjoint analyses is to treat the choice data that arise from the experiments like observed market choice data. In what follows, I present a discrete choice model that describes respondents' (stated) preferences for prices and star ratings. Each respondent $i = 1, ..., N$ receives a set of $m = 1, ..., M$ menus, each offering $j = 1, ..., J_m$ alternatives.[10] Each option j per menu m is characterized by two plan characteristics: prices, p_{jm}, and star ratings, r_{jm}. The utility a respondent i receives from an option j in menu m is given by:

$$u^s_{ijm} = \alpha^s_i p_{jm} + \beta^s_i r_{jm} + \epsilon^s_{ijm}, \tag{3.1}$$

where α^s_i, β^s_i are random coefficients that represent the respondent's stated preferences for prices and star ratings, respectively, and are distributed as

[10] In this experiment, $N = 618$, $M \in \{4, 8\}$, and $J_m = 2 \; \forall \; m \in M$.

follows,

$$\begin{pmatrix}\alpha_i^s\\ \beta_i^s\end{pmatrix} = \begin{cases}\begin{pmatrix}\alpha_{i,c}^s\\ \beta_{i,c}^s\end{pmatrix} \sim N\left(\begin{pmatrix}\overline{\alpha}_c^s\\ \overline{\beta}^s\end{pmatrix}, \begin{pmatrix}\sigma_{\alpha,c}^s & \rho^s\\ \rho^s & \sigma_\beta^s\end{pmatrix}\right) & \text{w/ prob. } \phi^s\\ \\ \begin{pmatrix}\alpha_{i,nc}^s\\ \beta_{i,nc}^s\end{pmatrix} \sim N\left(\begin{pmatrix}\overline{\alpha}_{nc}^s\\ 0\end{pmatrix}, \begin{pmatrix}\sigma_{\alpha,nc}^s & 0\\ 0 & 0\end{pmatrix}\right) & \text{w/ prob. } 1-\phi^s\end{cases}, \quad (3.2)$$

where ϕ^s is the probability that the respondent cares for star ratings. Lastly, I assume that ϵ_{ijm}^s follows the Type I Extreme Value distribution. With this set of assumptions, the choice probability of a respondent i choosing plan j in menu m is given by:

$$P_{ijm}^s(p_{jm}, r_{jm}|\Phi_i^s = \phi_i^s; \theta^s) = \begin{cases}\frac{exp(\alpha_i^s p_{jm}+\beta_i^s r_{jm})}{\sum_{k=1}^{J_m} exp(\alpha_i^s p_{km}+\beta_i^s r_{km})} & \text{if } \phi_i^s = 1\\ \frac{exp(\alpha_i^s p_{jm})}{\sum_{k=1}^{J_m} exp(\alpha_i^s p_{km})} & \text{if } \phi_i^s = 0\end{cases}, \quad (3.3)$$

where $\theta^s = (\overline{\alpha}_c^s, \overline{\alpha}_{nc}^s, \overline{\beta}^s, \sigma_{\alpha,c}^s, \sigma_{\alpha,nc}^s, \sigma_\beta^s, \rho^s, \phi^s)$ are the parameters to be estimated.

Define vector $d_n \equiv (d_1, .., d_m, .., d_M)$ to be the portfolio of choices for respondent i, where $d_m \equiv (d_{1m}, .., d_{jm}, ..., d_{Jm})$ is a vector that shows which plan j in menu m was chosen by the respondent, with

$$d_{jm} = \begin{cases}1 & \text{if } i \text{ chooses } j \text{ in menu } m\\ 0 & \text{o/w}\end{cases}. \quad (3.4)$$

Then, the probability that respondent i chooses portfolio $d = (d_1, ..., d_m, ..., d_M)$ is given by:

$$P_{id}^s(d|r, p; \theta^s) =$$
$$\sum_{\phi_i^s \in \{0,1\}} P(\Phi_i^s = \phi_i^s) \cdot \int_{\alpha_i^s, \beta_i^s} \prod_{m=1}^{M} \prod_{j=1}^{J_m} P_{ijm}(p_{jm}, r_{jm}|\Phi_i^s = \phi_i^s; \theta^s)^{d_{jm}} dF(\alpha_i^s, \beta_i^s), \quad (3.5)$$

where $r \equiv (r_1, .., r_m, .., r_M)$; $r_m = (r_{1m}, .., r_{jm}, .., r_{Jm})$, and $p \equiv (p_1, .., p_m, .., p_M)$; $p_m = (p_{1m}, ..., p_{Jm})$.

By providing respondents a set of $M = \{4, 8\}$ menus to choose a plan per menu, I generated a panel of choices per respondent that contributed to the identification of the groups that do not care about star ratings. Specifically, the percent of these groups was identified by the proportion of respondents who always chose the cheapest option across all the different menus they were presented. Identification of the rest of the coefficients was accomplished as in standard random coefficient models. An advantage of this experimental design is the random assignment of the different menus to different respondents, the random assignment of plan attributes to different menus, and the assumption that the two options in every menu were identical in every other dimension. The combination of those aspects secure the orthogonality conditions needed for identification of the parameters without requiring additional instrumental variables.

To estimate the parameters of interest, I used the Simulated Maximum Likelihood method. I estimated the parameters for both ex ante *aware* and *unaware* respondents. Table 3.11 presents the results from the two different specifications. Specification (1) assumes that respondents who do not care about star ratings have the same price preferences as those who somewhat care about star ratings, i.e., $\alpha^s_{i,nc} = \alpha^s_{i,c} | \beta^s_{i,c} = 0 \sim N(\overline{\alpha}^s_c + \rho \frac{\sigma_{\alpha^s_c}}{\sigma^s_\beta}(0 - \overline{\beta}^s), \sigma^{s2}_{\alpha,c}(1 - \rho^{s2}))$. Specification (2) assumes that price preferences between *caring* and *non-caring* respondents follow different distributions, but there is no correlation between the price and star rating random coefficients. The assumptions in these two specifications are reasonable and also reduce the number of parameters to be estimated given the number of observations in the survey. Estimation of all the parameters would require a much higher number of observations, which was beyond the scope of the study.

All estimates in specification (1) are statistically significant. The value of ϕ^s indicates that the proportion of respondents who care about star ratings is higher for the *aware* types in comparison to the *unaware*. Moreover, conditional on caring, the *aware* types place a value on an extra star rating of \$30.11 per month, while the *unaware* types place a value on an extra star rating of \$27.54 per month. Overall, the *aware* types value an extra star rating \$25.23 per month, while the *unaware* types value an extra star rating \$20.66 per month. This difference is not surprising since *aware* consumers tend to be more conscious when they make their choices as compared to *unaware* consumers. The estimates of specification (2) are close to those of specification (1). They are mostly statistically significant, with the exception of $\overline{\alpha}_{nc^s}$ and $\sigma^s_{\alpha,nc}$ for the *aware* types.

I will use the estimates from the first specification for further analysis, as they are fairly robust and also statistically significant. Overall, the results imply that consumers are willing to pay \$25 per month for a marginal increase in star ratings. The amount Medicare pays via the Quality Bonus Program to incentivize firms to increase star ratings is \$30 per month per plan on average. This observation further indicates that the estimates are reasonable.

Table 3.12 presents the results across these two specifications for the two roll-outs of the survey separately. Although the estimates might seem slightly different in magnitude, the overall value for star ratings are similar for all consumer types and all specifications between the two roll-outs.

Table 3.13 presents results across two more specifications: one in which I estimate all the parameters of the model without making any restrictions and one in which I assume that consumers who care about star ratings have different price preferences from those who do not care and also that price preferences of consumers who do not care about star ratings are homogenous. Overall, the results are similar to the results of the first two specifications. Although the model can be as flexible as the one of specification (3), such a specification requires a higher number of survey respondents, as currently some of the extra parameters cannot be estimated precisely. Table 3.14 also presents the results across these two specifications for the two roll-outs of the survey separately.

Lastly, tables 3.15, 3.16, 3.17, 3.18 present results of the main specifications of interest ((1) and (2)) assuming that demographic characteristics, age, health, income, and education, respectively, affect consumer preferences. Not particular patterns and/or particular conclusions can be made based on these results, as many of the additional variables tend to be statistically insignificant. For more precise estimates that can lead to better conclusions there is need for an additional round of surveys with a higher number of respondents overall.

Overall, the results across all different specifications suggest that consumer willingness to pay for star ratings amounts to \$20 – \$25 on average depending on their type. This result combined with the results of the previous chapter indicate that under the assumption of perfect information consumer willingness to pay for star ratings would be underestimated and would lead to biased welfare predictions. Further, it is worth to notice that the survey demand estimates predict that consumers are willing to pay on average 50% – 85% more than the logit demand estimates of the previous chapter predict. This observation support the validity of the survey as this level is close to the proportion of consumers who are unaware of the SRS.

3.1.5 Investigating the general process of a choice decision

In the last part of the survey, I include a set of questions that investigate the general process Medicare beneficiaries follow when they make their annual plan enrollment decision. First, I ask respondents whether they receive any help when they choose plans. In case they do, I further investigate what sources they use to receive help and also what kind of help they receive. Tables 3.19, 3.20, and 3.21 present the results. Overall, 20% of the Medicare population receive help when choosing a plan. It is interesting to notice that this level does not change depending on consumers' SRS awareness status. Among the beneficiaries who receive help, it seems that the most common source they use is either a family member or an independent agent who usually suggest or choose a plan for them.

Second, I ask respondents to rate the importance they place on different plan characteristics–premium, copay/ coinsurance, networks, star ratings, drug coverage, extra benefits (dental, vision, hearing)–on a 5-point scale–Not important, Slightly important, Moderately important, Important,Very important–when they make their annual plan enrollment decision. Figure 3.9 shows the question as it appeared to respondents and table 3.22 presents the results. The results indicate that the most important characteristic for consumers are networks and drug prescription coverage. Premiums and copays/coinsurance rates follow, while rest benefits (dental, vision, hearing) and star ratings are the least important for them. Importantly, although preferences for the different plan characteristics do not differ between aware and unaware consumers, they do differ with respect to star ratings. This observation validates the conjoint experiment results that suggest that the different consumer types have different levels of willingness to pay for star ratings.

Third, I ask respondents an open-ended question where I encourage them to report any additional factor they take into consideration before they make their annual enrollment decision.[11] The last question of this part of the survey asks respondents to report their satisfaction level with respect to MA and/or TM on a scale from 1-100. Table 3.23 presents the results. Overall, consumers tend to be slightly more satisfied with MA compared to TM, while the levels they report do not change regarding their SRS awareness status.

[11] See section 2.3 of the appendix for respondents' answers.

3.2 Conclusions

This chapter describes the survey I conducted to elicit information on the proportion of the Medicare population that was aware of the SRS and elicit consumer preferences for star ratings. In the survey, I also recover further information on the ways consumers choose plans and on the most important factors that affect their annual enrollment decisions.

The results indicated that a high percent of the population is unaware of the SRS implying that the information provision policy is neither informing consumers nor is it providing an incentive mechanism for competition in quality on the supply side. The results coming from the conjoint experiments suggest that had consumers been aware of the SRS, they would value star ratings, but not at the same level as the consumers who are already aware. This is not surprising; consumers who are more aware of any kind of information tend to be more conscious when they make their purchase decisions.

Using conjoint experiments to identify consumer preferences is a novel part of this study as it provides unique estimates for a group of people researchers do not have information on. The exogenous variation of prices and star ratings further supports demand estimates that are not contaminated by endogeneity issues that are prevalent when using real choice data.

A caveat of the conjoint analysis, however, is that the recovered random coefficient for star ratings is a composite of consumer preferences for star ratings and consumer value for the information the star ratings provide. Hence, I cannot separately identify the value consumers have for the information the SRS provides with respect to clinical outcomes and customer service from the exact preferences consumers have for the this specific dimension of quality per se. A potential solution to this problem is building a Bayesian learning discrete choice model that would exploit respondents' prior and posterior beliefs on star ratings. Building and estimating such a model potentially requires an additional round of surveys that would generate variation in respondents' prior beliefs with respect to the SRS. This task is left for future research.

Lastly, the survey provides evidence that although consumers value star ratings, they tend to value other plan characteristics more. Having insurance plans that provide their preferred provider networks along with drug coverage seems to be the characteristics for which consumers would be willing to pay the the most.

	Mean	Median	Std. Dev.
Total duration (in min)	9.186	7.800	8.198
Specific questions (in sec)			
Filtering	36.077	29.474	28.598
Awareness	17.525	14.138	16.423
HCE 1st	18.677	15.754	15.702
HCE 4th	7.928	6.433	6.000
HCE 8th	7.434	5.507	8.008
Demographics	53.900	47.303	27.004

Table 3.1: Survey duration

	1st roll-out			2nd roll-out		
	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
Total duration (in min)	9.693	8.216	9.624	8.135	7.250	4.127
Specific questions (in sec)						
Filtering	36.412	30.782	28.778	35.382	27.553	28.281
Awareness	16.784	13.864	12.227	19.063	15.102	22.748
HCE 1st	18.848	15.799	17.507	18.321	15.732	11.089
HCE 4th	7.814	6.412	5.352	8.163	6.552	7.168
HCE 8th	-	-	-	7.434	5.507	8.008
Demographics	54.790	48.734	26.819	52.054	44.049	27.358

Table 3.2: Survey duration split by roll-outs

	Survey		MCBS	
Variable	MA	TM	MA	TM
MA enrollment indicator	1	0	1	0
Demographics				
Spending ($)	24758	26074		
Income ($)			32484	34239
Age	70.372	70.376	73.13	72.3
Female	.480	.543	.561	554
Black	.028	.037	.087	.075
Hispanic	.006	0	.027	.014
Education				
High School	.298	.345	.341	.343
Bachelor degree	.337	.358	.148	.166
Attended college			.211	.199
Graduate degree	.201	.166		
Health Status				
Excellent	.092	.123	.162	.155
Very good	.432	.314	.297	.274
Good	.353	.370	.309	.305
Fair	.116	.166	.176	.180
Poor	.006	.018	.054	.082
Observations	456	162	17620	35647

Table 3.3: Demographics of the survey population

	1st roll-out		2nd roll-out	
Variable	MA	TM	MA	TM
MA enrollment indicator	1	0	1	0
Demographics				
Spending ($)	25970	28529	22203	21560
Age	71.757	71.962	67.462	67.375
Female	.482	.509	.476	.607
Black	.012	.028	.061	.053
Hispanic	.006	0	.006	0
Education				
High School	.304	.339	.285	.357
Bachelor degree	.317	.349	.380	.375
Graduate degree	.216	.188	.170	.125
Health Status				
Excellent	.097	.132	.081	.107
Very good	.475	.339	.340	.267
Good	.310	.386	.442	.339
Fair	.113	.113	.122	.267
Poor	.003	.018	.013	.017
Observations	309	106	147	56

Table 3.4: Demographics of the survey population split by roll-outs

Consumer types	Aware	Unaware
Population		
Traditional Medicare	0.173	0.827
Medicare Advantage	0.321	0.679
Overall	0.226	0.774

Table 3.5: *Aware*/*Unaware* consumer types

	1st roll-out		2nd roll-out	
Consumer types	Aware	Unaware	Aware	Unaware
Population				
Traditional Medicare	0.140	0.860	0.250	0.750
Medicare Advantage	0.322	0.678	0.327	0.673
Overall	0.180	0.820	0.220	0.780

Table 3.6: Awareness levels in the different roll-outs

	Medicare Advantage		Traditional Medicare	
Variable	Aware	Unaware	Aware	Unaware
Demographics				
Age	70.212	70.448	68.96	70.67
Female	.397	.519	.500	.552
Black	.027	.029	0	.044
Hispanic	.006	.006	0	0
Education				
High School	.246	.322	.285	.358
Bachelor's degree	.390	.312	.357	.358
Graduate degree	.226	.190	.250	.149
Health Status				
Excellent	.102	.087	.250	.097
Very good	.431	.432	.357	.305
Good	.342	.358	.250	.395
Fair	.109	.119	.142	.171
Poor	.013	.003	0	.022
Plan Characteristics				
Monthly premium	59.810	84.662		
Star rating	4.489			
Observations	146	310	28	134

Table 3.7: Awareness by MA/TM

	Medicare Advantage		Traditional Medicare	
Variable	Aware	Unaware	Aware	Unaware
Demographics				
Age	71.591	71.834	70.000	72.260
Female	.448	.497	.357	.532
Black	.030	.004	.000	.032
Hispanic	.000	.009	.000	.000
Education				
High School	.244	.331	.214	.358
Bachelor's degree	.397	.279	.357	.347
Graduate degree	.193	.227	.357	.163
Health Status				
Excellent	.132	.080	.357	.097
Very good	.479	.473	.357	.336
Good	.306	.312	.214	.413
Fair	.081	.127	.071	.119
Poor	.000	.004	.000	.021
Plan Characteristics				
Monthly premium	61.765	90.734		
Star rating	4.483			
Observations	98	211	14	92

Table 3.8: Awareness by MA/TM - 1st roll-out

	Medicare Advantage		Traditional Medicare	
Variable	Aware	Unaware	Aware	Unaware
Demographics				
Age	67.395	67.494	67.928	67.190
Female	.291	.565	.642	.595
Black	.020	.080	.000	.071
Hispanic	.020	.000	.000	.000
Education				
High School	.250	.303	.357	.357
Bachelor's degree	.375	.383	.357	.380
Graduate degree	.291	.111	.142	.119
Health Status				
Excellent	.041	.101	.142	.095
Very good	.333	.343	.357	.238
Good	.416	.454	.285	.357
Fair	.166	.101	.214	.285
Poor	.041	.000	.000	.023
Plan Characteristics				
Monthly premium	55.900	71.270		
Star rating	4.500			
Observations	48	99	14	42

Table 3.9: Awareness by MA/TM - 2nd roll-out

Attribute	Levels
Prices	0, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 25, 26, 27, 28, 29, 30, 31, 32, 33, 37, 39, 41, 42, 43, 44, 46, 49, 50, 51, 52, 53, 54, 56, 57, 60, 61, 62, 63, 64, 66, 67, 69, 70, 71, 72, 73, 75, 78, 80, 81, 82, 87, 88, 89, 99, 101
Star ratings	2, 2.5, 3, 3.5, 4, 4.5, 5

Table 3.10: Attributes for conjoint experiments

	(1)		(2)	
Consumer types	Aware	Unaware	Aware	Unaware
Variable				
$\overline{\alpha}_c^s$	-.099	-.120	-.085	-.104
	(.014)	(.010)	(.013)	(.009)
$\sigma_{\alpha,c}^s$	.064	.064	.046	.052
	(.013)	(.009)	(.009)	(.007)
$\overline{\beta}^s$	2.992	3.324	2.563	2.840
	(.414)	(.277)	(.363)	(.236)
σ_{β}^s	1.537	1.397	1.353	.958
	(.835)	(.558)	(.422)	(.283)
ρ^s	.012	.007		
	(.029)	(.017)		
ϕ^s	.837	.750	.880	.771
	(.117)	(.064)	(.091)	(.043)
$\overline{\alpha}_{nc}^s$	-.100	-.121	-.368	-.334
	(.059)	(.033)	(.550)	(.164)
$\sigma_{\alpha,nc}^s$	.064	.064	.205	.207
	(.066)	(.036)	(.404)	(.112)
Observations	177	447	177	447
Assumptions				
$\alpha_{i,nc}^s = \alpha_{i,c}^s \vert \beta_{i,c}^s = 0,\ \rho \neq 0$	Y	Y	N	N
$\alpha_{i,nc}^s \neq \alpha_{i,c}^s \vert \beta_{i,c}^s = 0,\ \rho = 0$	N	N	Y	Y
Monthly star value for *caring* types	\$30.114	\$27.547	\$30.157	\$27.221
Overall monthly star value	\$25.232	\$20.665	\$26.547	\$21.012

Table 3.11: Conjoint experiment estimates

	1st roll-out				2nd roll-out			
Specification	(1)	(1)	(2)	(2)	(1)	(1)	(2)	(2)
Consumer types	Aware	Unaware	Aware	Unaware	Aware	Unaware	Aware	Unaware
Variable								
$\overline{\alpha}_c^s$	-.082	-.0111	-.084	-.097	-.101	-.135	-.091	-.111
	(.015)	(.014)	(.014)	(.013)	(.019)	(.015)	(.008)	(.009)
$\sigma_{\alpha,c}^s$	.057	.050	.061	.056	.052	.067	.046	.063
	(.018)	(.018)	(.013)	(.010)	(.017)	(.012)	(.006)	(.008)
$\overline{\beta}^s$	2.334	3.094	2.372	2.642	3.102	4.223	2.750	3.233
	(.355)	(.319)	(.339)	(.332)	(.686)	(.468)	(.241)	(.113)
σ_β^s	.003	.813	.363	.564	2.127	1.931	1.995	.881
	(1.175)	(.705)	(.754)	(.510)	(.939)	(.890)	(.292)	(.115)
ρ^s	.023	.067			.014	.005		
	(.048)	(.024)			(.034)	(.030)		
ϕ^s	.926	.698	.928	.761	.763	.697	.843	.720
	(.048)	(.048)	(.064)	(.067)	(.140)	(.069)	(.046)	(.044)
$\overline{\alpha}_{nc}^s$	-9.261	-.124	-6.996	-.144	-.103	-.135	-.284	-.472
	(.057)	(.047)	(5.223)	(.071)	(.047)	(.038)	(.142)	(.064)
$\sigma_{\alpha,nc}^s$	.057	.050	.874	.078	.052	.067	.167	.270
	(.058)	(.042)	(9.501)	(.064)	(.045)	(.037)	(.073)	(.036)
Observations	115	306	115	306	62	141	62	141
Assumptions								
$\alpha_{i,nc}^s = \alpha_{i,c}^s \vert \beta_{i,c}^s = 0,\ \rho \neq 0$	Y	Y	N	N	Y	Y	N	N
$\alpha_{i,nc}^s \neq \alpha_{i,c}^s \vert \beta_{i,c}^s = 0,\ \rho = 0$	N	N	Y	Y	N	N	Y	Y
Monthly star value given $\Phi_i = 1$	$28.173	$27.809	$28.238	$27.237	$30.469	$31.297	$30.219	$29.126
Overall monthly star value	$26.097	$19.424	$26.204	$20.727	$23.257	$21.819	$25.475	$20.970

Table 3.12: Conjoint experiment estimates in two roll-outs

	(3)		(4)	
Consumer types	Aware	Unaware	Aware	Unaware
Variable				
$\overline{\alpha}^s_c$	-.083	-.114	-.077	.093
	(.011)	(.010)	(.007)	(.006)
$\sigma^s_{\alpha,c}$	.030	.044		
	(.024)	(.011)		
$\overline{\beta}^s$	2.487	2.824	2.433	2.383
	(.372)	(.267)	(.321)	(.230)
σ^s_β	1.224	1.366	1.700	2.006
	(.460)	(.471)	(.327)	(.220)
ρ^s	.017	.001		
	(.019)	(.019)		
ϕ^s	.875	.783	.887	.898
	(.091)	(.055)	(.074)	(.045)
$\overline{\alpha}^s_{nc}$	-.461	-.624	-.203	-.451
	(1.023)	(.550)	(.167)	(.462)
$\sigma^s_{\alpha,nc}$	.404	.764		
	(.891)	(.685)		
Observations	177	447	177	447
Assumptions				
$\alpha^s_{i,c} \neq \alpha^s_{i,nc}, \rho \neq 0$	Y	Y	N	N
$\sigma_{\alpha,\{c,nc\}} = 0, \overline{\alpha}^s_c \neq \overline{\alpha}^s_{nc}, \rho = 0$	N	N	Y	Y
Monthly star value given $\Phi_i = 1$	\$29.724	\$24.765	\$31.477	\$25.575
Overall monthly star value	\$26.010	\$19.400	\$27.936	\$22.967

Table 3.13: More conjoint experiment estimates

	1st roll-out				2nd roll-out			
Specification	(3)	(3)	(4)	(4)	(3)	(3)	(4)	(4)
Consumer types	Aware	Unaware	Aware	Unaware	Aware	Unaware	Aware	Unaware
Variable								
$\overline{\alpha}_c^s$	-.084	.-099	-.068	-.084	-.099	-.113	-.077	-.102
	(.015)	(.014)	(.010)	(.008)	(.026)	(.015)	(.012)	(.008)
$\sigma_{\alpha,c}^s$	.061	.049			.070	.067		
	(.043)	(.018)			(.020)	(.012)		
$\overline{\beta}^s$	2.372	2.740	2.340	2.086	2.790	3.177	2.884	2.979
	(.346)	(.341)	(.388)	(.400)	(.672)	(.366)	(.431)	(.320)
σ_β^s	.363	.464	1.175	1.758	1.665	1.216	1.953	2.187
	(.519)	(.704)	(.399)	(.330)	(.921)	(.548)	(.482)	(.286)
ρ^s	.050	.050			.020	.011		
	(.022)	(.027)			(.034)	(.026)		
ϕ^s	.928	.734	.863	.913	.817	.763	.787	.815
	(.061)	(.055)	(.106)	(.131)	(.125)	(.044)	(.081)	(.049)
$\overline{\alpha}_{nc}^s$	-.084	-.126	-.163	-.119	-.261	-2.569	-.165	-9.505
	(2.8e+)	(.044)	(.197)	(.110)	(.295)	(1.023)	(.072)	5.705
$\sigma_{\alpha,nc}^s$	.010	.059			.144	1.233		
	(1.4e+)	(.042)			(.244)	(.336)		
Observations	115	306	115	306	62	141	62	141
Assumptions								
$\alpha_{i,c}^s \neq \alpha_{i,nc}^s, \rho \neq 0$	Y	Y	N	N	Y	Y	N	N
$\sigma_{\alpha,\{c,nc\}} = 0, \overline{\alpha}_c^s \neq \overline{\alpha}_{nc}^s, \rho = 0$	N	N	Y	Y	N	N	Y	Y
Monthly star value given $\Phi_i = 1$	\$28.238	\$27.676	\$34.411	\$24.833	\$28.181	\$28.115	\$37.356	\$28.983
Overall monthly star value	\$26.204	\$20.314	\$29.697	\$22.672	\$23.024	\$21.451	\$29.476	\$23.621

Table 3.14: More conjoint experiment estimates in two roll-outs

Consumer types	(1) Aware	(1) Unaware	(2) Aware	(2) Unaware
Variable				
$\overline{\alpha}^s_{c,old}$	-.122	-.111	-.103	-.143
	(.074)	(.025)	(.155)	(.050)
$\overline{\alpha}^s_{c,young}$	.030	.004	.018	.039
	(.079)	(.020)	(.156)	(.050)
$\overline{\alpha}^s_{c,middle}$	.026	.006	.015	.043
	(.083)	(.027)	(.157)	(.053)
$\sigma^s_{\alpha,c,old}$	.061	.040	.047	.052
	(.014)	(.009)	(.010)	(.007)
$\overline{\beta}^s_{old}$	2.816	1.837	1.335	2.915
	(2.310)	(1.478)	(.432)	(1.265)
$\overline{\beta}^s_{young}$	.016	.007	-.019	-.079
	(2.610)	(1.459)	(4.733)	(1.278)
$\overline{\beta}^s_{middle}$	-.043	.012	.103	-.031
	(2.709)	(1.539)	(4.773)	(1.374)
$\sigma^s_{\beta_{old}}$	1.450	2.210	1.335	.912
	(.790)	(.246)	(.432)	(.284)
ρ^s	.010	.009		
	(.027)	(.010)		
ϕ^s	.855	.980	.880	.768
	(.136)	(.064)	(.091)	(.042)
$\overline{\alpha}^s_{nc}$	-.123	-.111	-.365	-.329
	(.223)	(.148)	(.537)	(.157)
$\sigma^s_{\alpha,nc}$	.061	.039	.203	.205
	(.072)	(.031)	(.392)	(.107)
Observations	177	447	177	447
Assumptions				
$\alpha^s_{i,nc} = \alpha^s_{i,c} \vert \beta^s_{i,c} = 0,\ \rho \neq 0$	Y	Y	N	N
$\alpha^s_{i,nc} \neq \alpha^s_{i,c} \vert \beta^s_{i,c} = 0,\ \rho = 0$	N	N	Y	Y
Monthly star value for *caring* types - young	30.814	17.147	30.321	27.101
Overall monthly star value - young	26.355	16.804	26.690	20.839
Monthly star value for *caring* types - middle	28.697	17.537	30.538	28.703
Overall monthly star value - middle	24.544	17.186	26.881	22.071
Monthly star value for *caring* types - old	22.952	16.435	24.967	20.283
Overall monthly star value - old	19.631	16.399	21.977	15.596

Table 3.15: Conjoint experiment estimates - Demographics: age

	(1)		(2)	
Consumer types	Aware	Unaware	Aware	Unaware
Variable				
$\overline{\alpha}^s_{c,poor}$	-.065	-.112	-.026	-.134
	(.053)	(.034)	(.779)	(.091)
$\overline{\alpha}^s_{c,excellent}$	-.057	.009	-.086	.033
	(.082)	(.037)	(.779)	(.094)
$\overline{\alpha}^s_{c,verygood}$	-.059	-.001	-.048	.027
	(.073)	(.032)	(.780)	(.091)
$\overline{\alpha}^s_{c,good}$	-.038	.015	-.067	.034
	(.072)	(.028)	(.780)	(.091)
$\overline{\alpha}^s_{c,fair}$	-.040	-.007	-.062	.016
	(.085)	(.034)	(.779)	(.093)
$\sigma^s_{\alpha,c,poor}$	.077	.038	.051	.052
	(.014)	(.009)	(.011)	(.007)
$\overline{\beta}^s_{poor}$	2.958	1.815	2.459	2.894
	(2.658)	(1.025)	(2.635)	(2.187)
$\overline{\beta}^s_{excellent}$	-.429	.027	.081	.063
	(2.895)	(1.124)	(2.631)	(2.260)
$\overline{\beta}^s_{verygood}$	.792	-.004	.153	-.367
	(2.863)	(1.040)	(2.633)	(2.201)
$\overline{\beta}^s_{good}$	.066	-.018	.004	.072
	(2.838)	(.986)	(2.634)	(2.201)
$\overline{\beta}^s_{fair}$	-.168	...	-.006	.278
	(2.944)	(1.088)	(2.630)	(2.245)
$\sigma^s_{\beta_{poor}}$	1.868	2.163	1.280	.997
	(.860)	(.249)	(.343)	(.281)
ρ^s	.020	.010		
	(.037)	(.010)		
ϕ^s	.848	.976	.895	.784
	(.132)	(0.059)	(.047)	(.043)
$\overline{\alpha}^s_{nc}$	-.067	-.112	-1.730	-.358
	(.093)	(.089)	(8.646)	(.191)
$\sigma^s_{\alpha,nc}$	.077	.038	.911	.223
	(.089)	(.011)	(4.750)	(.132)
Observations	177	447	177	447
Assumptions				
$\alpha^s_{i,nc} = \alpha^s_{i,c} \mid \beta^s_{i,c} = 0,\, \rho \neq 0$	Y	Y	N	N
$\alpha^s_{i,nc} \neq \alpha^s_{i,c} \mid \beta^s_{i,c} = 0,\, \rho = 0$	N	N	Y	Y
Monthly star value for *caring* types - poor	45.481	16.134	93.580	21.550
Overall monthly star value - poor	38.567	15.746	83.765	16.915
Monthly star value for *caring* types - fair	26.494	14.936	27.849	26.910
Overall monthly star value - fair	22.466	14.577	24.928	21.121
Monthly star value for *caring* types - good	29.341	18.449	26.307	29.780
Overall monthly star value - good	22.880	18.043	23.548	23.374
Monthly star value for *caring* types - very good	30.060	15.895	35.053	23.695
Overall monthly star value - very good	25.490	15.545	31.376	18.598
Monthly star value for *caring* types - excellent	20.685	17.801	22.541	29.284
Overall monthly star value - excellent	17.540	18.201	20.176	22.985

Table 3.16: Conjoint experiment estimates - Demographics: health

	(1)		(2)	
Consumer types	Aware	Unaware	Aware	Unaware
Variable				
$\overline{\alpha}^s_{c,above25}$	-.216	-.160	-.068	-.089
	(.025)	(.014)	(.018)	(.014)
$\overline{\alpha}^s_{c,below18}$	.148	.029	-.031	-.036
	(.029)	(.012)	(.023)	(.018)
$\overline{\alpha}^s_{c,above18_below25}$	.057	.051	-.016	-.014
	(.025)	(.014)	(.025)	(.019)
$\sigma^s_{\alpha,c,above25}$	.107	.071	.050	.041
	(.016)	(.009)	(.009)	(.014)
$\overline{\beta}^s_{above25}$	3.080	2.377	2.474	2.779
	(.328)	(.275)	(.528)	(.413)
$\overline{\beta}^s_{below18}$	-.740	-.041	.012	.011
	(.052)	(.033)	(.673)	(.495)
$\overline{\beta}^s_{above18_below25}$	.270	-.115	-.002	.026
	(.067)	(.037)	9.678)	(.531)
$\sigma^s_{\beta_{above25}}$	1.905	2.870	1.336	1.167
	(.505)	(.327)	(.426)	(.241)
ρ^s	.035	.010		
	(.017)	(.013)		
ϕ^s	.960	.950	.904	.769
	(.045)	(.043)	.081	(.042)
$\overline{\alpha}^s_{nc}$	-.210	-.160	-.449	-.573
	(.189)	(.092)	(.862)	(.330)
$\sigma^s_{\alpha,nc}$	.107	.071	.252	.565
	(.114)	(.052)	(.648)	(.303)
Observations	177	447	177	447
Assumptions				
$\alpha^s_{i,nc} = \alpha^s_{i,c} \vert \beta^s_{i,c} = 0,\ \rho \neq 0$	Y	Y	N	N
$\alpha^s_{i,nc} \neq \alpha^s_{i,c} \vert \beta^s_{i,c} = 0,\ \rho = 0$	N	N	Y	Y
Monthly star value for *caring* types - > 25	14.228	14.796	36.038	30.965
Overall monthly star value - > 25	13.658	14.056	32.603	23.839
Monthly star value for *caring* types - ≤ 18	34.173	17.819	24.870	22.060
Overall monthly star value - ≤ 18	32.806	16.928	22.499	16.983
Monthly star value for *caring* types - $> 18 \leq 25$	21.022	20.648	29.009	27.031
Overall monthly star value - $> 18 \leq 25$	20.181	19.615	26.243	20.810

Table 3.17: Conjoint experiment estimates - Demographics: income

	(1)		(2)	
Consumer types	Aware	Unaware	Aware	Unaware
Variable				
$\overline{\alpha}^s_{c,other}$	-.185	-.134	-.123	-.109
	(.025)	(.015)	(.033)	(.019)
$\overline{\alpha}^s_{c,highschool}$	.105	.017	.037	-.003
	(.025)	(.013)	(.038)	(.022)
$\overline{\alpha}^s_{c,bachelor}$	.109	.032	.039	.014
	(.022)	(.013)	(.035)	(.023)
$\overline{\alpha}^s_{c,graduate}$	.117	.039	.061	.014
	(.026)	(.015)	(.037)	(.023)
$\sigma^s_{\alpha,c,other}$	.070	.035	.044	.050
	(.011)	(.010)	(.010)	(.007)
$\overline{\beta}^s_{other}$	3.013	1.832	2.626	2.825
	(.364)	(.232)	(.827)	(.553)
$\overline{\beta}^s_{highschool}$	.067	-.019	.064	-.018
	(.100)	(.049)	(1.052)	(.617)
$\overline{\beta}^s_{bachelor}$	-.273	-.040	-.025	.014
	(.074)	(.050)	(.933)	(.635)
$\overline{\beta}^s_{graduate}$	-.036	.024	-.358	-.011
	(.102)	(.067)	(.956)	(.696)
$\sigma^s_{\beta_{other}}$	.707	2.314	1.292	.955
	(.589)	(.264)	9.435)	(.286)
ρ^s	.035	0.011		
	(.027)	(.010)		
ϕ^s	.833	0.973	.874	.769
	(.047)	(.057)	(.067)	(.042)
$\overline{\alpha}^s_{nc}$	-.190	-.133	-.646	-.343
	(.173)	(.863)	(1.186)	(.168)
$\sigma^s_{\alpha,nc}$	.070	.035	.412	.215
	(.125)	(.024)	(.805)	(.118)
Observations	177	447	177	447
Assumptions				
$\alpha^s_{i,nc} = \alpha^s_{i,c} \vert \beta^s_{i,c} = 0,\, \rho \neq 0$	Y	Y	N	N
$\alpha^s_{i,nc} \neq \alpha^s_{i,c} \vert \beta^s_{i,c} = 0,\, \rho = 0$	N	N	Y	Y
Monthly star value for *caring* types - other	16.273	13.661	21.274	25.737
Overall monthly star value - other	13.565	13.292	18.613	19.814
Monthly star value for *caring* types - high school	38.849	17.543	31.460	24.753
Overall monthly star value - high school	32.384	17.069	27.526	19.057
Monthly star value for *caring* types - bachelor	36.207	17.624	30.950	29.690
Overall monthly star value - bachelor	30.182	17.148	27.079	22.857
Monthly star value for *caring* types - graduate	44.209	19.521	36.774	29.631
Overall monthly star value - graduate	36.852	18.993	32.175	22.812

Table 3.18: Conjoint experiment estimates - Demographics: education

Use help?	Yes	No
Consumer types		
Aware	.201	.798
Unaware	.207	.792
Overall	.205	.794

Table 3.19: Do consumers use help?

Sources	
Counselor from SHIP	.039
Medicare helpline	.062
Family member	.377
Independent agent/broker	.440
Personal insurance company	.047
Other	.035

Table 3.20: Sources where consumers receive help when choosing plans

Types of help consumers receive	
They suggested/chose a plan	.669
They showed me PlanFinder	.094
They helped me use PlanFinder	.086
Other	.151

Table 3.21: Ways consumers receive help when choosing plans

Variable	Not	Slightly	Moderately	Important	Very
Premium					
Aware	.028	.063	.183	.281	.442
Unaware	.027	.056	.159	.270	.486
Copay/ Coinsurance					
Aware	.022	.040	.166	.379	.390
Unaware	.027	.033	.119	.391	.427
Networks					
Aware	.028	.011	.091	.316	.551
Unaware	.029	.024	.110	.313	.522
Star Ratings					
Aware	.034	.074	.224	.465	.201
Unaware	.175	.114	.337	.247	.123
Drug Prescription					
Aware	.051	.057	.086	.293	.511
Unaware	.065	.049	.094	.277	.513
Benefits					
Aware	.068	.126	.235	.264	.304
Unaware	.105	.083	.220	.272	.317

Table 3.22: Ranking different plan characteristics

	MA	TM
Consumer types		
Aware	87.993	83.392
Unaware	85.761	83.082
Overall	86.475	83.135

Table 3.23: MA/TM ranking

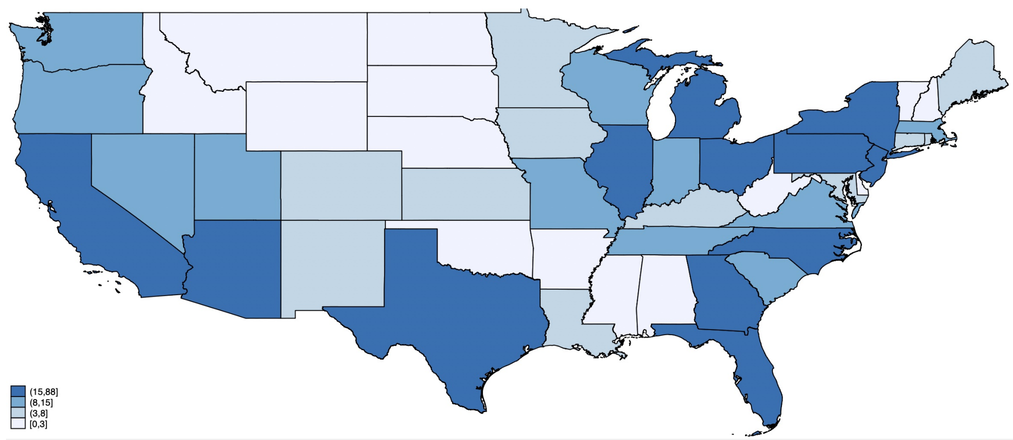


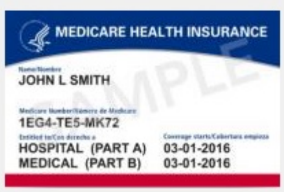
Figure 3.1: Geographic variation of the survey population

Medicare-eligible individuals have two main ways of getting hospital and medical coverage. These are Original Medicare and Medicare Advantage. The insurance card(s) you show when you go to the doctor likely indicate whether you have chosen Original Medicare or Medicare Advantage.

Below are example health insurance cards. Please look at them carefully, read through their coverage descriptions and choose the option that currently applies to you.

Note: *The cards shown are just for illustration. Your actual cards will be containing your own personal information.*

(Please select one option below)



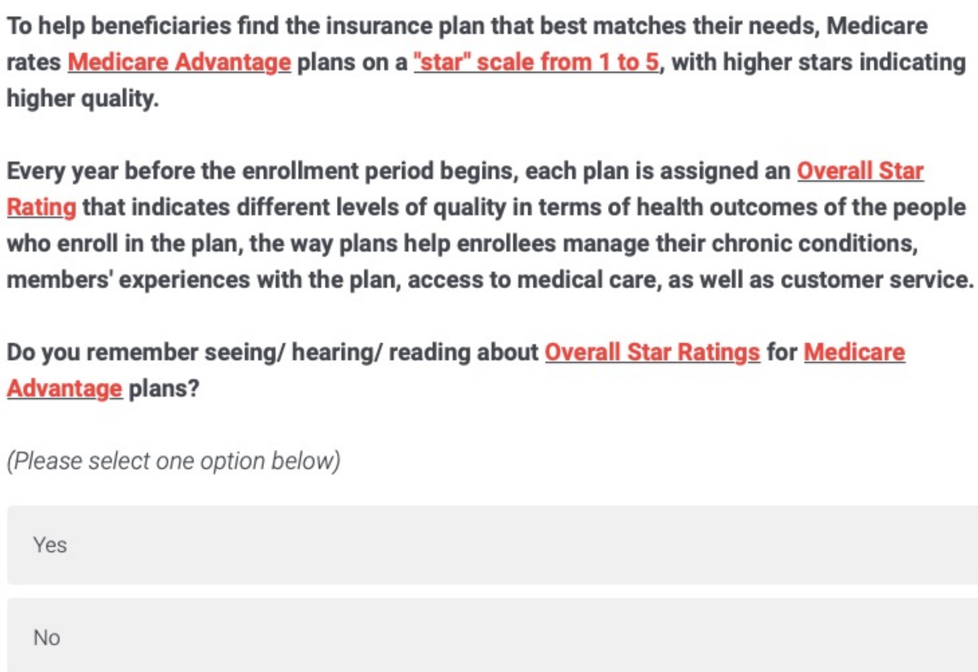
Original Medicare: is directly provided by the government and includes **hospital and medical** coverage.

If you are in Original Medicare, you may also have **supplemental coverage** (also known as **Medigap**) which you receive through a private insurer. In this case, you carry an extra card.

Medicare Advantage: is provided through private health insurance companies, like Aetna, Humana, Blue Cross Blue Shield, etc. These companies cover everything Original Medicare covers. Sometimes, they also offer dental, hearing, and vision benefits.

Note: *The cards shown above are* ***only examples*** *of private insurance companies. Your actual insurer might be different.*

Figure 3.2: Filtering respondents



To help beneficiaries find the insurance plan that best matches their needs, Medicare rates Medicare Advantage plans on a "star" scale from 1 to 5, with higher stars indicating higher quality.

Every year before the enrollment period begins, each plan is assigned an Overall Star Rating that indicates different levels of quality in terms of health outcomes of the people who enroll in the plan, the way plans help enrollees manage their chronic conditions, members' experiences with the plan, access to medical care, as well as customer service.

Do you remember seeing/ hearing/ reading about Overall Star Ratings for Medicare Advantage plans?

(Please select one option below)

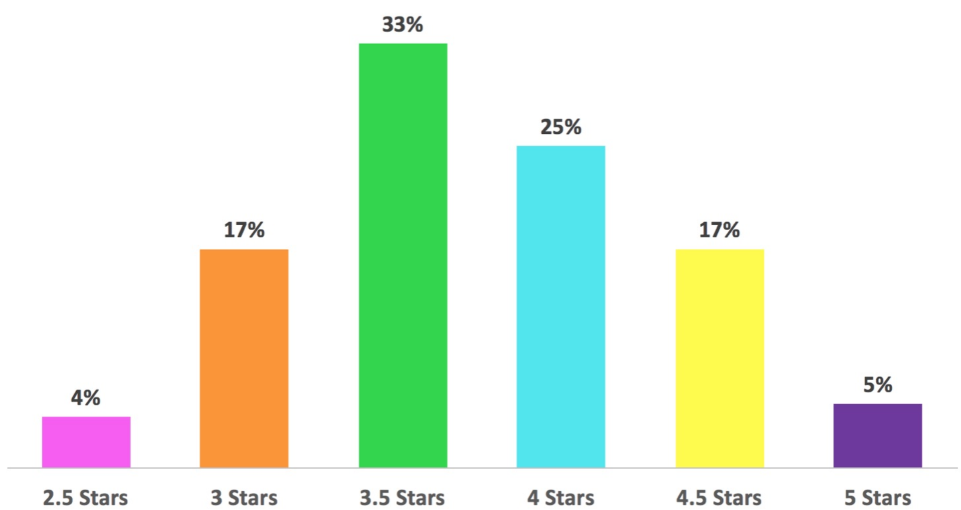
Yes

No

Figure 3.3: Eliciting information on information.

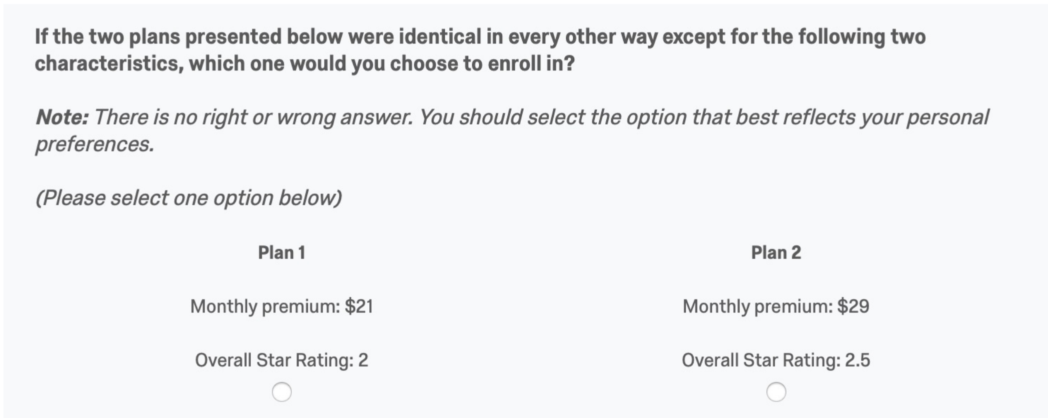
Medicare Advantage plans can take different forms, each offering different options. In 2019, the number of plans offered in the area you live was on average 10, and the typical plan received a 3.5 Overall Star Rating.

The chart below shows the percentage of Medicare Advantage plans that received certain levels of Overall Star Ratings.



Note: *Remember that an Overall Star Rating indicates different levels of quality in terms of health outcomes of the people who enroll in the plan, the way plans help enrollees manage their chronic conditions, members' experience with the plan, access to medical care, as well as customer service. More stars indicate higher quality.*

Figure 3.4: Informing respondents on the star ratings distribution.



If the two plans presented below were identical in every other way except for the following two characteristics, which one would you choose to enroll in?

Note: *There is no right or wrong answer. You should select the option that best reflects your personal preferences.*

(Please select one option below)

Plan 1	**Plan 2**
Monthly premium: $21	Monthly premium: $29
Overall Star Rating: 2	Overall Star Rating: 2.5

Figure 3.5: HCE example question

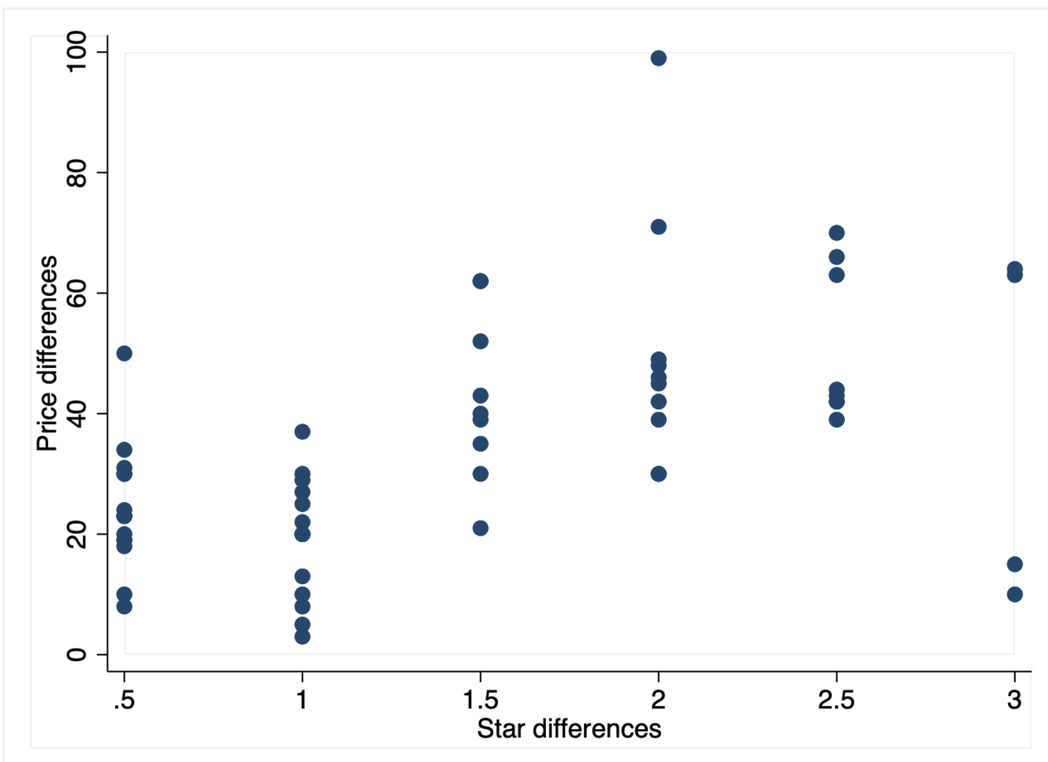


Figure 3.6: Price - Star differences variation in the conjoint experiments

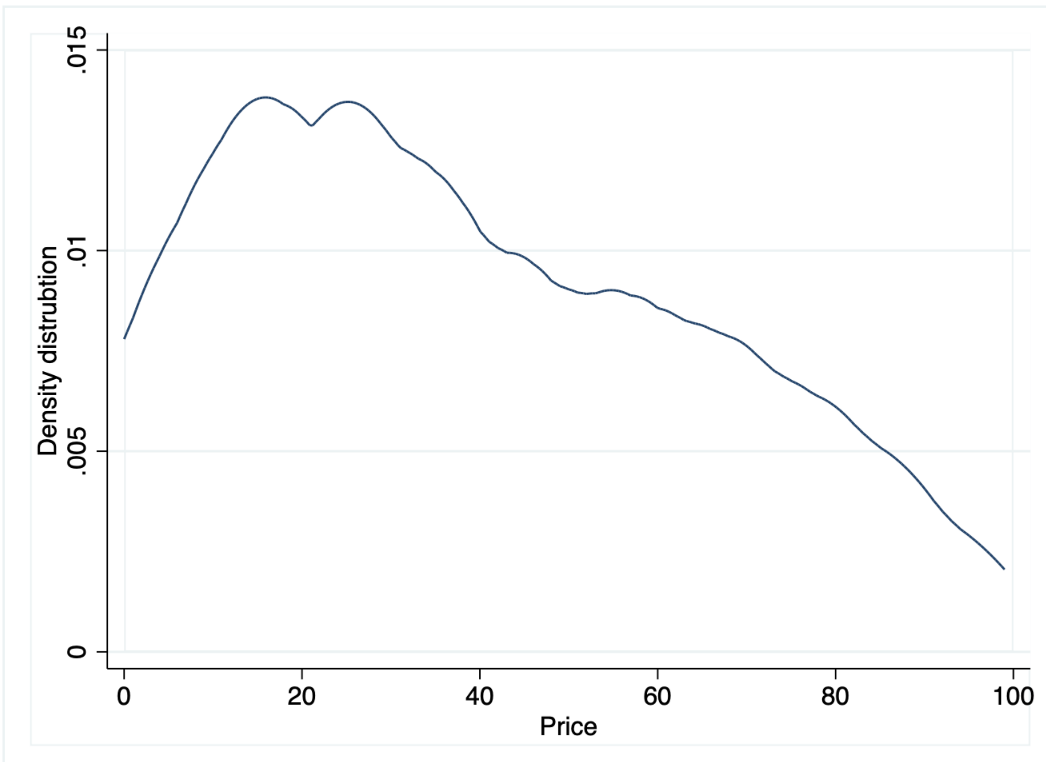


Figure 3.7: Price variation in the conjoint experiments

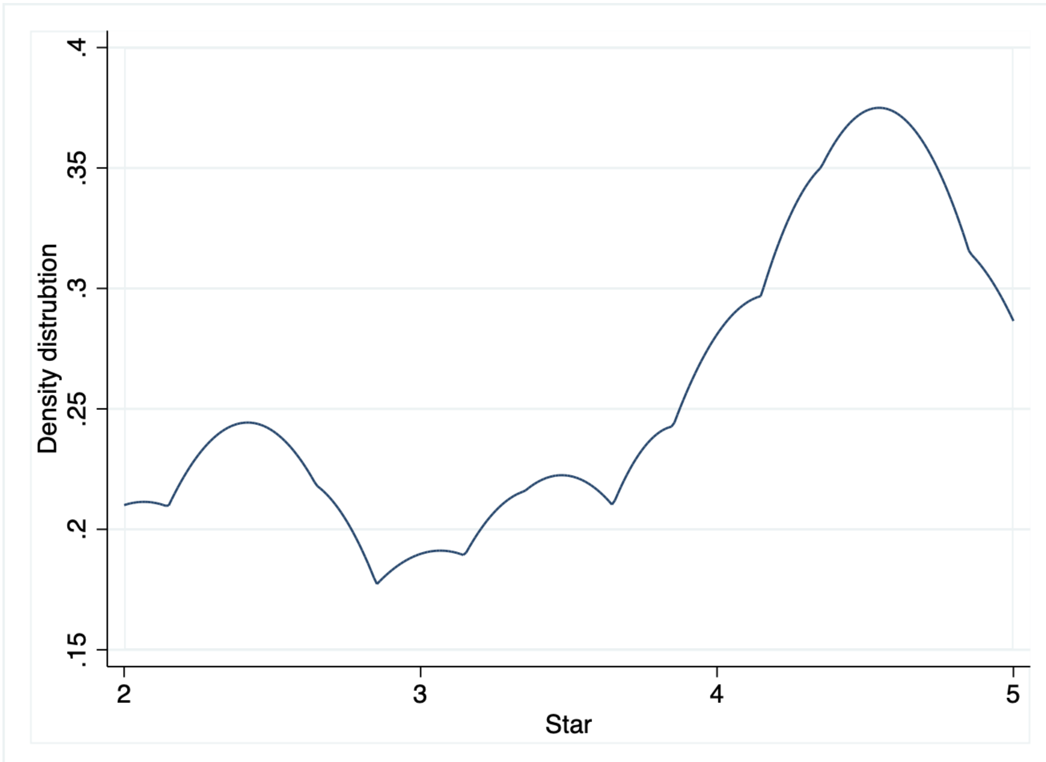


Figure 3.8: Star variation in the conjoint experiments

Please indicate how important each of the following plan characteristic was when you made your annual health insurance choice for 2019.

(Please select an "importance" level for each plan characteristic below)

	Not Important	Slightly Important	Moderately Important	Important	Very Important
Monthly premium	○	○	○	○	○
Copay/ Coinsurance	○	○	○	○	○
Networks of providers in the plan	○	○	○	○	○
Overall Star Ratings	○	○	○	○	○
Drug Coverage	○	○	○	○	○
Extra benefits (dental, vision, hearing)	○	○	○	○	○

Figure 3.9: Survey question referring to importance levels of plan characteristics in consumer plan choices

Chapter 4

Structural Model

In this chapter, I build and estimate a full demand and supply structural equilibrium model. On the demand side I build a flexible model that takes into account the different existing consumer types–aware and unaware of the Star Rating System (SRS)–combining the publicly available data that are provided by the Centers for Medicare and Medicaid Services (CMS) with my survey data. On the supply side I model insurers to compete on both prices and star ratings, while also allowing for uncertainty to govern their choice for star ratings. Finally, I use the estimates to conduct welfare analysis investigating a series of different counterfactual scenarios.

4.1 Demand for Health Insurance Plans

In this section, I introduce a Bayesian learning discrete choice model in the spirit of [26]. My model further allows for consumer awareness heterogeneity with respect to the SRS. The model builds upon the demand model of [91] and the more recent models of [29] and [73].

4.1.1 Assumptions/ Primitives

As it is standard in the literature, I define markets at a county level, and I assume that, in every market, $m = 1, 2, ..., M$, there is a set of insurers, $k_m = 1, ..., K_m$, each offering a set of plans, $j_{km} = 1, ..., J_{km}$ in year t. I classify the different characteristics that determine the quality of a health insurance plan in three categories. The first category includes characteristics, x_{kjmt}, which refer to the general benefit design (e.g., drug, dental coverage), which is observed both by the consumers and the econometrician. The second category includes characteristics, ξ_{kjmt}, such as advertising and provider networks that are observed by consumers only.[1] The third category includes characteristics, q_{kjmt}, which refer to customer satisfaction and clinical outcomes. I assume that these characteristics are unobserved by both the consumers and the econometrician until the introduction of the Star Rating System (SRS). Plans in every market are differentiated with respect to premium, p_{kjmt}, and all other characteristics, $x_{kmjt}, \xi_{kjmt},$ and q_{kjmt}.

There are two main periods: the pre-SRS and the post-SRS period. In the pre-SRS period, consumers form some prior beliefs about q_{kjmt} and then choose their most preferred plan. In the post-SRS period, signals for quality, q_{kjmt}, are given to consumers in the form of star ratings, r_{kt}. Consumers who are aware of the SRS update their prior beliefs in a Bayesian fashion and then

[1]It is arguably accepted in the literature that advertising and provider networks are two characteristics that consumers observe and value when they make their choice decisions. With the exception of [6], researchers/econometricians can rarely observe this information.

choose their most preferred plan. Consumers who are not aware of the SRS behave as if they were in the pre-SRS period and choose their most preferred plan based on their prior beliefs.

Let u_{ikjmt} denote consumer i's utility when enrolled in plan j offered by insurer k in market m of year t. I assume that the consumer is risk neutral. Then, the expected utility she gains from this plan conditional on the plan characteristics and her personal taste is given by:

$$\begin{aligned} &E[u_{ikjmt}|\overline{x}_{kjmt}, \theta_i, \epsilon_{ikjmt}] = \\ &= \alpha_i p_{kjmt} + \beta_i E[q_{kjmt}|\overline{x}_{kjmt}, \theta_i, \epsilon_{ikjmt}] + \gamma x_{kjmt} + \xi_{kjmt} + \tilde{\epsilon}_{ikjmt}, \end{aligned} \tag{4.1}$$

where $\overline{x}_{kjmt} = (p_{kjmt}, x_{kjmt}, \xi_{kjmt})$ is a vector of the plan characteristics observed by the consumer, and $\theta_i = (\alpha_i, \beta_i)'$ is a vector that represents her preferences with respect to p_{kjmt} and q_{kjmt}, respectively. Finally, $\tilde{\epsilon}_{ikjmt}$ is a zero mean i.i.d. stochastic error term that follows the Type I extreme-value distribution across plans and consumers and represents i's idiosyncratic tastes. With this set of assumptions, caring for prices and other characteristics, the consumer chooses the plan j that maximizes her current expected utility.

I assume that consumers have heterogenous preferences for p_{kjmt} and q_{kjmt}, allowing for flexible substitution patterns that are not contaminated by the Independence of Irrelevant Alternatives property and also allowing for the possibility that consumers might not care at all about q_{kjmt}. Consumer preferences for p_{kjmt} and q_{kjmt} are then represented by the following distribution:

$$\theta_i = \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{cases} \begin{pmatrix} \alpha_{i,c} \\ \beta_{i,c} \end{pmatrix} \sim N\left(\begin{pmatrix} \overline{\alpha}_c \\ \overline{\beta} \end{pmatrix}, \begin{pmatrix} \sigma_{\alpha,c} & \rho \\ \rho & \sigma_\beta \end{pmatrix} \right) & \text{w/ prob. } \phi \\ \\ \begin{pmatrix} \alpha_{i,nc} \\ \beta_{i,nc} \end{pmatrix} \sim N\left(\begin{pmatrix} \overline{\alpha}_{nc} \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\alpha,nc} & 0 \\ 0 & 0 \end{pmatrix} \right) & \text{w/ prob. } 1-\phi \end{cases}, \tag{4.2}$$

where with probability ϕ consumer i cares about q_{kjmt} and with probability $1 - \phi$ she does not care about it. *Caring* consumers' preferences for p_{kjmt} and q_{kjmt} are represented by a vector of two random coefficients, $(\alpha_{i,c}, \beta_{ic})'$, which are jointly normally distributed according to the top part of eq. 4.2. *Non-caring* consumers' preferences for p_{kjmt} and q_{kjmt} are represented by a random coefficient, $\alpha_{i,nc}$, which is normally distributed, and a mass point at zero, respectively, according to the bottom part of eq. 4.2. One could possibly argue that consumers might not care about other characteristics as well. While in theory this might be true, in practice it has been shown that consumers care about both characteristics x_{kjmt} and ξ_{kjmt}.[2]

Consumers who do not choose a Medicare Advantage (MA) plan have the option to choose Traditional Medicare (TM), which is considered to be the outside option in this market. The utility consumer i receives from choosing TM is given by:

$$u_{i0mt} = \tilde{q}_0 + \beta_i q_{TM} + \tilde{\epsilon}_{i0mt}, \tag{4.3}$$

[2] Allowing for a more flexible model in which consumers have no preferences for other characteristics than q_{kjmt} is possible, but it is beyond the scope of this study.

where q_{TM} represents the quality of TM with respect to clinical outcomes and customer service.[3] I normalize the utility gained from the outside option by subtracting u_{i0mt} from each u_{ikjmt}. Then, the utility specification of eq. 4.1 takes the following form:

$$E[u_{ikjmt}|\overline{x}_{kjmt}, \theta_i, \epsilon_{ikjmt}] =$$

$$= q_0 + \alpha_i p_{kjmt} + \beta_i(E[q_{kjmt}|\overline{x}_{kjmt}, \theta_i, \epsilon_{ikjmt}] - q_{TM}) + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt}, \tag{4.4}$$

where $\epsilon_{ikjmt} = \tilde{\epsilon}_{ikjmt} - \tilde{\epsilon}_{i0mt}$ is distributed logistic as the difference of two type I extreme value distributions and $q_0 = -\tilde{q}_0$.

4.1.2 Choice of Health Insurance Plan

In this section, I describe the demand model before and after the introduction of the SRS. The utility specification depends on the different existing consumer types, as in [13].

4.1.2.1 Choice of Health Insurance Plan Before SRS

Although plan characteristics x_{kjmt} and prices p_{kjmt} could affect consumer beliefs about q_{kjmt}, I make the simplifying assumption that consumers do not use this information to form their beliefs. Characteristics x_{kjmt} that compose the benefit design of a plan do not reflect either customer satisfaction or clinical outcomes that compose quality, q_{kjmt}, a fact that justifies this

[3] This is the same dimension of quality the SRS is providing information on.

assumption.[4] The positive relation between star ratings and prices shown on table 2.2 would make it more reasonable to assume that prices contain some information about this dimension of quality. For computational tractability, I assume that consumers do not take this relation into account. However, on the supply side I will allow firms to price star ratings. Hence, consumers perceive that quality, q_{kjmt}, is drawn from the following distribution:

$$q_{kjmt} \sim N(\mu_0, \sigma_0^2), \tag{4.5}$$

and the mean of their prior beliefs is given by:

$$E[q_{kjmt}|p_{kjmt}, x_{kjmt}, \xi_{kjmt}, \theta_i, \epsilon_{ikjmt}] = \mu_0. \tag{4.6}$$

Under this assumption and conditional on prices and all the dimensions of quality, consumer i chooses the plan that maximizes her current expected utility, which takes the following form:

$$E[u_{ikjmt}|\overline{x}_{kjmt}, \theta_i, \epsilon_{ikjmt}] = q_0 + \alpha_i p_{kjmt} + \beta_i(\mu_0 - q_{TM}) + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt}. \tag{4.7}$$

The set of individual attributes that lead to the choice of plan j is defined as follows:

$$\begin{aligned} &A_{kjmt}(x_{.mt}, p_{.mt}, \xi_{.mt}; \theta) = \\ &= (\theta_i, \epsilon_{i0mt}, ..., \epsilon_{ikjmt} | u_{ikjmt} \geq u_{iljmt} \forall l = 0, 1, ..., K, j = 1, 2, ..., J_k). \end{aligned} \tag{4.8}$$

[4] The fact that there is no relation between star ratings–which are the signals of q_{kjmt}–and plan characteristics, as table 2.2 shows, further validates this point.

Given the distributional assumption on ϵ_{ikjmt}, the market share of j-th plan is given by:

$$s_{kjmt}(x_{.mt}, p_{.mt}, \xi_{.mt}; \theta) =$$
$$= \int\limits_{\theta_i} \frac{exp(q_0 + \alpha_i p_{kjmt} + \beta_i(\mu_0 - q_{TM}) + \gamma x_{kjmt} + \xi_{kjmt})}{1 + \sum\limits_{k \in K} \sum\limits_{\substack{j' \in J \\ j' \neq j}} exp(q_0 + \alpha_i p_{kj'mt} + \beta_i(\mu_0 - q_{TM}) + \gamma x_{kj'mt} + \xi_{kj'mt})} d\theta_i f(\theta_i), \tag{4.9}$$

where $\theta_i = (\alpha_i, \beta_i)'$ is distributed as a mixture of two normal distributions with density $f(\theta_i) = \phi f(\theta_{i,c}) + (1-\phi) f(\theta_{i,nc})$, with $\theta_{i,c} = (\alpha_{i,c}, \beta_{i,c})'$ and $\theta_{i,nc} = (\alpha_{i,nc}, \beta_{i,nc})'$. The expected consumer surplus for an individual i given prices and all dimensions of quality is given by:

$$E(CS_i) = \frac{1}{\alpha_i} ln(1 + \sum_j exp(q_0 + \alpha_i p_{kjmt} + \beta_i(\mu_0 - q_{TM}) + \gamma x_{kjmt} + \xi_{kjmt})). \tag{4.10}$$

4.1.2.2 Choice of Health Insurance Plan After SRS

After the introduction of the SRS, individuals receive signals for quality, q_{kjmt}, in the form of star ratings, r_{kjmt}. Star ratings are assigned at a contract level, and all plans offered by a specific contract get the same star rating. Remember that every insurer for every market in which it operates signs a contract with Medicare under which the insurer offers a set of different plans, each corresponding to a different benefit design. Hence, by assigning star ratings at a contract level, CMS assigns star ratings at an insurer-market, km, level. Since all plans j offered by an insurer k get the same star ratings, for the rest of the analysis I will suppress the index j in variables q_{kjmt} and r_{kjmt}. I assume that the signal through the star rating is distributed normally around the true quality, q_{kmt} as follows:

$$r_{kmt} \sim N(q_{kmt}, \sigma_r^2), \tag{4.11}$$

where σ_r represents the noise the star ratings send to consumers. Consumers who are aware of the SRS receive these signals and update their beliefs in a Bayesian fashion. The mean of their posterior beliefs becomes:

$$E[q_{kmt}|p_{kjmt}, x_{kjmt}, \xi_{kjmt}, r_{kmt}, \theta_i, \epsilon_{ikjmt}] = wr_{kmt} + (1-w)\mu_0, \tag{4.12}$$

where w represents the weight consumers put on the signal they receive and is defined by $w = \frac{\sigma_0^2}{\sigma_0^2+\sigma_r^2}$. If the star ratings are precise, then $\sigma_r \to 0$ and $w \to 1$, and thus when forming their posterior beliefs, consumers put all the weight on the signals. Conversely, if the star ratings are very noisy, then $\sigma_r \to \infty$ and $w \to 0$, and thus when forming their posterior beliefs, consumers place no weight on the signals.[5] Consumers who are not aware of the SRS behave as if they were in the pre-SRS period.

Let Λ_i be a discrete random variable that indicates whether consumer i is aware of the SRS and follows the Bernoulli distribution, $\Lambda_i \sim Bern(\lambda)$, as follows:

$$\Lambda_i = \begin{cases} 1 & \text{w/ prob. } \lambda \\ 0 & \text{w/ prob. } 1-\lambda \end{cases}. \tag{4.13}$$

[5]Under the assumption that both the prior beliefs and the signal are normally distributed, the variance of the posterior beliefs is $Var[q_{kt}] = w\sigma_r^2$. However, since I have assumed that consumers are risk neutral, I cannot identify it.

Under these assumptions and conditional on all the plan characteristics and her type, consumer i will choose the plan that maximizes her current expected utility, which takes the following form:

$$E[u_{ikjmt}|\overline{x}_{kjmt}, r_{kmt}, \theta_i, \epsilon_{ikjmt}, \lambda_i] =$$
$$= \begin{cases} q_0 + \alpha_i p_{kjmt} + \beta_i(wr_{kmt} + (1-w)\mu_0 - q_{TM}) + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \Lambda_i = 1 \\ q_0 + \alpha_i p_{kjmt} + \beta_i(\mu_0 - q_{TM}) + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \Lambda_i = 0 \end{cases}. \tag{4.14}$$

Given the distributional assumption on ϵ_{ikjmt}, the market share of the j-th plan is given by:

$$s_{kjmt}(x_{.mt}, p_{.mt}, r_{.t}, \xi_{.mt}; \theta) =$$
$$= \lambda \int_{\theta_i} \frac{exp(q_0 + \alpha_i p_{kjmt} + \beta_i(wr_{kmt} + (1-w)\mu_0 - q_{TM}) + \gamma x_{kjmt} + \xi_{kjmt})}{1 + \sum_{k \in K} \sum_{\substack{j' \in J \\ j' \neq j}} exp(q_0 + \alpha_i p_{kj'mt} + \beta_i(wr_{kmt} + (1-w)\mu_0 - q_{TM}) + \gamma x_{kj'mt} + \xi_{kj'mt})} d\theta_i f(\theta_i)$$
$$+ (1-\lambda) \int_{\theta_i} \frac{exp(q_0 + \alpha_i p_{kjmt} + \beta_i(\mu_0 - q_{TM}) + \gamma x_{kjmt} + \xi_{kjmt})}{1 + \sum_{k \in K} \sum_{\substack{j' \in J \\ j' \neq j}} exp(q_0 + \alpha_i p_{kj'mt} + \beta_i(\mu_0 - r_{TM}) + \gamma x_{kj'mt} + \xi_{kj'mt})} d\theta_i f(\theta_i), \tag{4.15}$$

where $\theta_i = (\alpha_i, \beta_i)'$ is distributed as a mixture of two normal distributions, as noted in the previous section. The expected consumer surplus for an individual i given prices and all dimensions of quality in the post-SRS period is given by:

$$E(CS_i) = \lambda \frac{1}{\alpha_i} ln(1 + \sum_j exp(q_0 + \alpha_i p_{kjmt} + \beta_i(wr_{kmt} + (1-w)\mu_0 - q_{TM}) + \gamma x_{kjmt} + \xi_{kjmt}))+$$
$$+ (1-\lambda) \frac{1}{\alpha_i} ln(1 + \sum_j exp(q_0 + \alpha_i p_{kjmt} + \beta_i(\mu_0 - q_{TM}) + \gamma x_{kjmt} + \xi_{kjmt})). \tag{4.16}$$

The set of parameters to be estimated is $(q_0, q_{TM}, \overline{\alpha}_c, \overline{\beta}, \sigma_{\alpha,c}, \rho, \overline{\alpha}_{nc}, \sigma_{\alpha,nc}, \phi, \lambda, \gamma, w, \mu_0)$.

As mentioned in the second chapter, a major challenge arises in separately identifying consumers who are not aware of the SRS from consumers who do not care about star ratings, r_{kt}, since all these different consumer types seem observationally equal in the aggregate level data. Intuitively, if we assume that the star ratings are perfectly informative, i.e., $w = 1$, consumers who are not aware of the SRS behave as if they do not care about quality, q_{kt}. Taking into account that the star ratings might not be perfect signals of q_{kmt} complicates the situation further.[6]

I overcome this challenge, using the data that are coming from the survey I conducted. Having identified the groups of consumers who are *aware*/ *unaware* of the SRS, and having elicited both groups' preferences for star ratings, I combine these results into the main demand model, and I recover the rest parameters of interest. It is important to notice, here, that respondents' preferences for star ratings represent a composite of consumer preferences for q_{kmt} and the weight, w, they put on the information the star ratings provide as signals of q_{kmt}.[7]

4.1.3 Estimation

After I recover from the survey the proportions of the population that are *aware* and *unaware* of the SRS as well as the preferences of each group for

[6]See section 1.2 of the appendix for further details on the identification challenge in this model specifically.

[7]An additional round of surveys in which I "change" respondents' prior beliefs about quality would be necessary to recover the weight, w consumers put on the information they receive.

star ratings, I estimate the parameters of the main demand model following the algorithm used by [76]. In doing so, I plug the estimated random coefficients, $\hat{\alpha}_i^s, \hat{\beta}_i^s$, which I recovered from the survey, into the main demand model. Because the coefficient $\hat{\beta}_i^s$ refers to preferences for star ratings, I assume that star ratings are perfect signals of quality, q_{kjmt}, in the main model, and thus the weight consumers put on the information they receive amounts to one, i.e., $w = 1$. Combining two different sources of choice data–revealed preference and stated preference–requires that I control for potential differences in the idiosyncratic tastes between survey respondents and actual consumers. To do so, I use a rescale parameter, κ. Rescaling has the advantage of preserving the shape of the distribution of the preferences recovered from the survey in the main model, while at the same time it leads to more consistent estimates. ([90], [10], [19], [40]) The utility specification is given by:

$$u_{ikjmt} = \\ = \begin{cases} q_0 + \kappa \cdot (\hat{\alpha}_i^s p_{kjmt} + \hat{\beta}_i^s (r_{kt} - q_{TM})) + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{w/ prob. } \hat{\lambda} \\ q_0 + \kappa \cdot (\hat{\alpha}_i^s p_{kjmt} + \hat{\beta}_i^s (\mu_0 - q_{TM})) + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{w/ prob. } 1 - \hat{\lambda} \end{cases}, \tag{4.17}$$

which, given the distributional assumption of ϵ_{ikjmt}, results in the following market shares equation:

$$s_{kjmt}(x_{.mt}, p_{.mt}, r_{.mt}, \xi_{.mt}; \theta) =$$
$$= \hat{\lambda} \int\limits_{\hat{\alpha}_i^s, \hat{\beta}_i^s} \frac{exp(q_0 + \kappa(\hat{\alpha}_i^s p_{kjmt} + \hat{\beta}_i^s (r_{kt} - q_{TM})) + \gamma x_{kjmt} + \xi_{kjmt})}{1 + \sum\limits_{k \in K} \sum\limits_{\substack{j' \in J \\ j' \neq j}} exp(q_0 + \kappa(\hat{\alpha}_i^s p_{kj'mt} + \hat{\beta}_i^s (r_{kt} - q_{TM})) + \gamma x_{kj'mt} + \xi_{kj'mt})} dF(\hat{\alpha}_i^s, \hat{\beta}_i^s) +$$
$$+ (1 - \hat{\lambda}) \int\limits_{\hat{\alpha}_i^s, \hat{\beta}_i^s} \frac{exp(q_0 + \kappa(\hat{\alpha}_i^s p_{kjmt} + \hat{\beta}_i^s (\mu_0 - q_{TM})) + \gamma x_{kjmt} + \xi_{kjmt})}{1 + \sum\limits_{k \in K} \sum\limits_{\substack{j' \in J \\ j' \neq j}} exp(q_0 + \kappa(\hat{\alpha}_i^s p_{kj'mt} + \hat{\beta}_i^s (\mu_0 - q_{TM})) + \gamma x_{kj'mt} + \xi_{kj'mt})} dF(\hat{\alpha}_i^s, \hat{\beta}_i^s). \tag{4.18}$$

The set of parameters that needs to be estimated then reduces to $(q_0, q_{TM}, \mu_0, \gamma, \kappa)$.

The key point of the estimation exploits a population moment condition that is a product of instrumental variables and a structural error term, $\xi(\theta)$, to form a non-linear Generalized-Method-of-Moment (GMM) estimator, as in [14]. The key identifying assumption is the population moment condition that requires a set of exogenous instrumental variables. Formally, letting $Z = [z_1, ..., z_M]$ be a set of instruments, the required population moment condition is

$$E[Z \cdot \xi(\theta^*)] = 0, \tag{4.19}$$

where ξ is the error term as a function of the model parameters and θ^* denotes the true value parameters.

A common identification challenge is that differences in prices may reflect differences in the part of the quality, ξ_{kjmt}, that is unobserved by the econometrician. As mentioned in a previous section, the most important component of this part of quality is the plan's network of providers. To solve this problem, I rely on the panel structure of my data and exploit variation within an insurer contract in a similar fashion to [29]. This has the benefit, as the authors also mention, of utilizing the structure of the program under which plans within a contract share the same provider network. Moreover, I also use a set of time fixed effects to account for the possibility that the introduction of the QBP affected demand. Hence, I decompose ξ_{kjmt} as follows:

$$\xi_{kjmt} = \xi_k + \xi_t + \Delta\xi_{kjmt}, \tag{4.20}$$

where $\Delta\xi_{kjmt}$ measures the deviation in "market-specific" product tastes from these means.

The source of endogeneity is now limited to the set of unobservables that vary across the unobservables, $\Delta\xi_{kjmt}$, and the population moment condition is reduced to:

$$E[Z \cdot \Delta\xi(\theta^*)] = 0. \tag{4.21}$$

The GMM estimate is then given by:

$$\hat{\theta} = arg \min_{\theta} \Delta\xi(\theta)' Z A^{-1} Z' \Delta\xi(\theta), \tag{4.22}$$

where A is a consistent estimate of $E[Z'\Delta\xi\Delta\xi' Z]$. To compute these unobserved characteristics, $\Delta\xi_{kjmt}$, I first solve for the structural errors, ξ_{kjmt}, that solve the system of equations

$$s_{.mt}(x_{.mt}, p_{.mt}, r_{.mt}, \xi_{.mt}; \theta) = S_{.mt}, \tag{4.23}$$

where $s_{.mt}$ and $S_{.mt}$ are the predicted and observed market shares, respectively. Then, following [76] I regress the estimated fixed effects, $\hat{\xi}_{kjmt}$, on contract and time fixed effects, ξ_k, ξ_t, respectively, and compute the residual, $\hat{\Delta\xi}_{kjmt}$. Thus, the GMM estimate is finally defined by:

$$\hat{\theta} = arg \min_{\theta} \hat{\Delta}\xi(\theta)' Z A^{-1} Z' \hat{\Delta}\xi(\theta). \tag{4.24}$$

For the remaining source of endogeneity hidden in $\Delta\xi$, I use as instruments the average non-price attributes of other plans offered by the same insurer and the average non-price attributes of plans offered by competing insurers under the assumption that they are exogenous. These instruments are commonly known in the Industrial Organization literature as BLP instruments and are based on pricing concepts. Specifically, the idea is that each insurer will price each of its plans in a way that takes into account the substitution with other insurers' plans. For example, when an insurer is considering to increasing the price of its plan, consumers who will switch away from this plan to another of the same firm's plans do not represent as much of a loss as consumers who will switch to other firms' plans.

4.1.4 Results

Table 4.1 presents the estimates of Equation 4.17 across different specifications. All specifications assume that $r_{tm} = 4$ and that $\mu_0 = \overline{r}_{kt}$. The first two columns present estimates that do not include instrumental variables, while the last two present estimates after including instrumental variables.[8] The rescale parameter is slightly lower in the specifications that do not include instrumental variables compared to the ones that include them. The preference parameters for the different benefits are similar across all the specifications. Consistent with OLS estimates on price elasticities being closer to 0, I focus on the IV specifications (3) and (4). Specifically, my preferred specification is (4), as it includes both contract and year fixed effects.

Among all benefits, consumers are more willing to pay for drug coverage. Specifically, consumer willingness to pay for drug coverage is on average $170. Such a high level is not surprising, as Part D coverage is very costly for beneficiaries. [36] find a similar result. Consumer willingness to pay for the other characteristics is significantly lower. It is interesting to notice that consumer willingness to pay for star ratings based on the survey estimates is close to consumer willingness to pay for these benefits, and it is not as high as it is for drug coverage. This observation indicates that consumers do not value star ratings as much as they value drug coverage. However, whether the value for star ratings represents the actual value for the dimension of quality that refers to clinical outcomes and customer satisfaction is not clear. The mean own-price elasticity in my preferred specification amounts to -1.226, implying that a 1% increase in price will decrease the enrollment of a plan by 1.16%. This implies that consumers are moderately elastic, which is not surprising since MA is a subsidized market where most prices are close to zero.

[8]Standard errors are currently work in progress.

4.2 Supply Side

In this section, I introduce a model in which insurers compete on both prices, p, and quality, q, endogenizing the information environment and the financial incentives they receive. By allowing firms to compete on more than one dimension, the model resembles that of [44] and the more recent model of [73].

4.2.1 Assumptions/ Primitives

I allow competition to occur at the county level, which is standard in the literature ([91], [29], [73]). In each year, t, and every market, m, there is a total number of insurers, K_m, each indexed by $k = 1, 2, ..., K_m$. Every insurer offers a set of plans J_k, each indexed by $j = 1, 2, ..., J_k$. Insurers in every market and every year compete with each other simultaneously in two stages. In the first stage, they choose an average level of quality, $\tilde{q}_k$, which will characterize the set of plans they offer. In the second stage, star ratings, r_k, are realized, and then insurers choose prices, p_{kj}, for each plan they offer. For ease of exposition, in the rest of the section I suppress the indices mt.

The fact that the modes of the average summary rates distributions of figure 2.4 are not precisely at the thresholds where star ratings are assigned indicates that insurers cannot perfectly predict the final level of quality that will arise after they make their choice. CMS tends to make numerous changes in the algorithm it uses to construct the star ratings, some of which are pre-announced while others are not.[9] Hence, the final realized level of quality, q_k, is governed by some noise, η_k, that represents the uncertainty firms carry with respect to CMS' changes and is given by:

$$q_k = \tilde{q}_k + \eta_k, \tag{4.25}$$

where $\eta \sim N(0, \sigma_\eta^2)$. It is critical to remember that, although firms choose $\tilde{q}_k$, consumers observe only the final level of star rating, r_k, which results based on the following formula:

$$r_k = g(q_k) = \begin{cases} ... & \\ 2 & \text{if } 1.75 \leq q_k < 2.25 \\ ... & \\ 5 & \text{if } q_k \geq 4.75 \end{cases}. \tag{4.26}$$

4.2.2 Profit Function

The profit function that is relevant to insurer k's first-stage decision is given by:

$$\pi_k^I = \int\limits_{\eta_k, \eta_{-k}} \left(\pi_k^{II}(p^*(\tilde{q}, \eta), \tilde{q}, \eta; \chi)\right) dF(\eta_k, \eta_{-k}), \tag{4.27}$$

where $\pi_k^{II}(p^*(\tilde{q}, \eta), \tilde{q}; \chi)$ represents variable profits from offering insurance services given the equilibrium price, $p^*(\tilde{q}, \eta)$, which arises in the second stage. The variable profits, π_k^{II}, of the insurer are determined by the price, p_j, the

[9]CMS tends to make numerous changes regarding many aspects of the MA market, and it is the responsibility of insurers to keep track of them. It is common that insurers cannot do this perfectly.

insurer sets, the marginal cost, mc_j it incurs for each of the services it offers, and the subsidy, K_j, it receives for each plan j it offers, and they are given by:

$$\pi_k^{II}(p, r, x, \xi, \theta; \chi) = \sum_{j \in J_k} \left[(p_j - mc_j + K_j) \, M s_j(p, r, x, \xi; \theta) \right]. \tag{4.28}$$

I model the marginal cost, mc_j, to be log-linear in star ratings, r_j, characteristics, x_j, and some unobserved to the econometrician component, ω_j, as follows:

$$ln(mc_j) = \tau_r r_j + \tau_x x_j + \omega_j, \tag{4.29}$$

where $r_j = \frac{1}{J_k} r_k$.[10] The subsidy each plan receives per enrollee it serves is a function of the star rating and is given by:

$$K_j = K(\psi_t(r_k), \tilde{K}_j) = \psi_t(r_k) \cdot \tilde{K}_j, \tag{4.30}$$

where $\psi_t(r_k)$ represents the level of the bonus insurers receive depending on their star ratings for different years $t = \{2012, 2013, 2014, 2015, 2016\}$. For example, for $t = \{2015, 2016\}$, $\psi_t(r_k)$ is given by:

$$\psi_t(r_k) = \begin{cases} 1.05 & \text{if } r_k \geq 4 \\ 1 & \text{o/w} \end{cases}. \tag{4.31}$$

The formula is adjusted accordingly for years $t = \{2012, 2013, 2014\}$. Finally, $\tilde{K}_j$ represents the payment each plan receives from the government for each enrollee it serves after risk adjustment.

[10]By setting $r_j = \frac{1}{J_k} r_k$, I split the marginal cost a firm incurs among all the plans it offers equally.

4.2.3 Necessary Equilibrium Conditions

I now derive the optimality conditions for prices and quality.[11] These conditions will be used to identify the cost health insurers incur for the services they provide and the quality they produce.

An insurer makes two decisions in two sequential stages. I solve the insurer's problem backwards and derive the necessary equilibrium conditions for prices and quality as follows. Given the quality level, q_k, that arises in the first stage, the price optimality condition that arises in the second stage for each plan j insurer k offers is given by:

$$\frac{\partial \pi_k^{II}}{\partial p_j} = 0 \Rightarrow s_j(p,r) + \sum_{g \in J_k} (p_g - mc_g + K_g) \frac{\partial s_g(p,r)}{\partial p_j} = 0. \tag{4.32}$$

Using matrix notation, the First Order Conditions (FOCs) for the plans offered in a market are represented by:

$$s(p,r) - \Omega(p - mc + K) = 0, \tag{4.33}$$

where $\Omega = \Omega^*_{jl} * S_{jl}$, with $\Omega_{jl} = \begin{cases} 1 & \text{if } j,l \in J_k \\ 0 & \text{otherwise} \end{cases}$, and S_{jl} a matrix whose jl element is given by $s_{jl} = -\frac{\partial s_l(p,r)}{\partial p_j}$. The difference between this first-order condition and a standard first-order condition lies in the term K, which represents the subsidy each health insurer receives from Medicare.

In the first stage, insurers choose quality, $\tilde{q}_k$. Given the price equilibrium vector, $p^* = p(r)$ that arises in the second stage of the game, the problem

[11] Following the literature, I assume that a pure-strategy Nash equilibrium exists. Finding the conditions for the existence of the equilibrium is beyond the scope of this study.

each firm solves for each plan it offers is given by:

$$\pi_k^I = \sum_{z \in R_s} \left(\pi_k^{II}(p^*(z), z; x) \right) Prob(r(\tilde{q}, \eta) = z | \tilde{q}) \tag{4.34}$$

where z is a variable of integration that can take values $r \in R_s$, where $R_s = \{0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5\}$ represent all the values that star ratings can take. The FOCs of firm k w.r.t. $\tilde{q}_k$ will be:

$$\frac{\partial \pi_k^I}{\partial \tilde{q}_k} = 0 \Rightarrow \sum_{z \in R_s} \left(\pi_k^{II}(p^*(z), z; x) \right) \frac{\partial Prob(r(\tilde{q}, \eta) = z | \tilde{q})}{\partial \tilde{q}_k} = 0 \tag{4.35}$$

4.2.3.1 Estimation

I estimate the parameters of interest, τ, σ_η, in two stages. First, I estimate the marginal cost parameters, τ, using the FOCs with respect to price that arise in the second stage of the game. Second, given the estimated parameters, $\hat{\tau}$, I estimate the parameter σ_η using the Simulated Maximum Likelihood method. This parameter is identified by the spread of the modes around the thresholds at which star ratings are realized.

More precisely, in the first stage, I use the FOCs to recover the marginal cost firm k incurs for each plan j it offers, and then I use a simple OLS regression in the following model:

$$ln(p + K - \Omega^{-1} s(p, r)) = \tau \tilde{x} + \omega, \tag{4.36}$$

where $\tau = (\tau_r, \tau_x)$ is a vector of the marginal cost parameters and $\tilde{x} = (r, x)$ is a vector of star ratings and the rest of the plan characteristics.

In the second stage, the equation for estimation is represented by a simple linear model

$$q_k = \tilde{q}_k^* + \sigma_\eta \tilde{\eta}_k, \tag{4.37}$$

which is invertible in the error term, $\tilde{\eta}_k \sim N(0,1)$. Hence, I estimate the parameter σ_η by maximizing the following likelihood function, as follows:

$$\begin{aligned}
&\max_\sigma P(q_1, q_2, ..., q_K | \tilde{q}_1^*, \tilde{q}_2^*, ..., \tilde{q}_K^*; \sigma_\eta) = \\
&\max_\sigma \prod_{k=1}^{K} P(q_k | \tilde{q}_k^*; \sigma_\eta) = \\
&\max_\sigma \frac{1}{K} \sum_{k=1}^{K} ln(P(q_k | \tilde{q}_k; \sigma_\eta)) = \\
&\max_\sigma \frac{1}{K} \sum_{k=1} ln \left(\phi \left(\frac{q_k - \tilde{q}_k^*}{\sigma_\eta} \right) \cdot \frac{1}{\sigma_\eta} \right),
\end{aligned} \tag{4.38}$$

where $\tilde{q}_k^*$ is the equilibrium quality that arises after the firm solves its profit optimization problem in the first stage. The second line in the above equation results from the assumption that the errors η_k are independent across firms in every market. I estimate the above parameters using the following nested algorithm.

Given guesses for the parameters σ_η,

1. For each firm k and each plan j, solve for the prices p_{kj} that maximize the profits of the firm for each different star rating level.

2. Given the prices $p_{kj}^*(r)$ that arose in the previous step, for each firm k solve for quality $\tilde{q}_k$ that maximizes eq. 4.34

3. Given the optimized values $\tilde{q}_k^*$, solve for the parameters that maximize the likelihood function given in eq. 4.38.

4.2.4 Results

Table 4.2 reports estimates of equation 4.29. I include market fixed effects to account for costs that stem from geography. Specifications (2) and (4) also consider the possibility that the cost is increasing in star ratings, but at a decreasing rate. All the cost sharing plan design parameters enter with the correct sign and are statistically significant. My preferred specification and the one I use for the estimation of σ_η is specification (3). According to that specification, on average, the cost of a plan given all the benefits it offers is \$620. An extra star rating increases the cost of the firm by 4%. Hence, it costs the average firm \$25 to increase quality by an extra star rating. Offering drug coverage increases the cost of a firm by 21%, which is significantly higher than the cost the star ratings impose on insurers.

Lastly, the coefficient that captures the noise of the star ratings is given in Table 4.3. The standard deviation of the noise of star ratings amounts to 1.4 (in the range of 0 to 5), with a standard error equal to 0.035. Such a high standard deviation implies that, even though firms may target specific quality levels, the final star rating level that might be realized can be up to three standard deviations above or below the chosen level. This high magnitude is not surprising. The number of different individual metrics that are included in the composition of the final overall star rating is very high ($30-44$ performance metrics), and the model does not rationalize the variation that is coming from them.

4.2.5 Model Fit

To evaluate the model fit, I compare the predicted distributions of multiple variables that arise in equilibrium to the distributions of the corresponding variables that I observe in the data. Specifically, I focus on plan prices, government payments, and market shares. To construct the predicted distribution of prices, given the demand and supply estimates, I firstly compute the quality $\tilde{q}_k^*$ each firm chooses. Given this quality level, I calculate the probability that each star rating level will arise, $Prob(r(\tilde{q}, \eta) = z|\tilde{q})$. Having also calculated the prices that arise in equilibrium, $p^*(z; \theta, \chi)$ for every star rating level, I then calculate the expected price. I follow a similar procedure for the government expenditures and market shares. Figure 4.1 presents a kernel density plot of the predicted and observed distributions of the plan prices. The vertical line represents the mean of the respective distribution. Overall, across all observations, the average monthly price is observed to be \$51.67, while it is predicted to be \$38.11. Figures 4.2 and 4.3 present kernel density plots of the predicted and observed distributions of the government payments and firm market shares. The average monthly payment to a plan is observed to be \$833.03 and is predicted to be \$818.64, and the market share of a plan is observed to be 0.15 and is predicted to be 0.16.

4.3 Welfare analysis

In this section, I evaluate the relative impact of the two policies–SRS and QBP–on Consumer Surplus (CS), firm profits, government expenditures, and, ultimately, welfare. I compare outcomes that arise under three different levels of information structure: (i) full information, $\lambda = 1$, (ii) current level of information, $\lambda = 0.22$, and (iii) no information, $\lambda = 0$ and under two "supply-side" policy environments: (i) no quality subsidies provision and (ii) quality subsidies provision. Consequently, I end up with a combination of six different hypothetical environments that I compare and analyze.

For each set of counterfactuals, I compute the new quality, price, firm cost, firm profits, consumer surplus, and government expenditures that arise in equilibrium. In doing so, I use the estimates I recovered from my full equilibrium model of supply and demand. Given the demand and cost estimates in column (4) of Table 4.1 and column (3) of Table 4.2, respectively, I solve for the prices that arise in equilibrium for each star rating level using eq. 4.32. Given these prices and the estimate of Table 4.3, I then solve for the optimal level of quality, q_j^*, using eq. 4.35. Of course, the level of quality that will finally arise cannot be predicted since it depends on the distribution of the error η_k and the draw that will be realized. Hence, I proceed by computing the expected cost of a firm for providing its services, $E[C_j]$, as well as the expected price, $E[p_j^*]$, market shares, $E[s_j]$, profits, $E[\pi_j]$, consumer surplus, $E[CS_i]$ and government expenditures, $E[K_j]$, which arise in equilibrium. To keep computational tractability, I conduct these counterfactuals on monopolistic markets only.

4.3.1 Welfare outcomes

Table 4.4 summarizes outcomes under each of the six policy environments in a median monopolistic market. Table 4.7 also presents the equilibrium outcomes that arise on average taking into account all monopolistic markets. None of the patterns that are observed in table 4.4 change on average. Hence, I focus on one monopolistic market only to better describe the intuition behind the outcomes that arise. The table shows the level of quality chosen in equilibrium, $\tilde{q}_j^*$, the firm's expected profits, $E[\pi_j]$, the expected consumer surplus, $E[CS_i]$, and the expected payment, $E[K_j]$, the plan receives from Medicare. The top panel presents the outcomes that arise without the provision of quality subsidies under the three different information environments of interest. Similarly, the bottom panel presents the outcomes that arise with the provision of quality subsidies under the three different information environments of interest. All monetary values are expressed in dollars per month per plan per enrollee.

As indicated in the table, quality, $\tilde{q}_j^*$, increases as consumers become more informed under both "supply-side" policy environments. Figure 4.4 provides a better sense of the impact of each policy on the quality that is realized in equilibrium. The graph presents the equilibrium quality levels that arise for different levels of information under both "supply-side" regulatory environments. The horizontal axis represents the different levels of information, λ, and the vertical axis represents the levels of quality, $\tilde{q}_j^*$, that are chosen by the firm in equilibrium. Clearly, both the information provision and the quality bonus subsidy policies increase quality. Interestingly, it turns that, if 50% of the Medicare population was aware of the SRS, the market could realize the same level of quality under the implementation of the "demand-side" policy only as it currently does at $\lambda = 0.22$ under the implementation of both the "demand- and supply-side" policies.

Table 4.4 further shows that in more informed environments the profitability of the firm decreases on average. This is not surprising. Firms are able to experience higher profits when consumers are not well informed about the true quality by charging higher prices. Tables 4.8, 4.9, and 4.10 explain this by displaying the equilibrium prices, market shares, and profits for each star rating level, respectively, under all the hypothetical scenarios of interest. In the case where consumers are uninformed, i.e., $\lambda = 0$, at star rating levels that are lower than consumers' prior beliefs about quality, $r_j \leq \mu_0 = 3.7$, firms set high prices, taking advantage of consumers' inaccurate beliefs about the actual level of quality. Similarly, because consumers believe that these plans are better than what they actually are, the demand is higher compared to the case where consumers are aware of the exact level of quality. Consequently, firms are able to experience higher profits when consumers are uninformed. The opposite patterns are observed when the star rating levels are above the level of consumers' prior beliefs, $r_j > \mu_0 = 3.7$. Interestingly, comparing the profits that arise under the different "supply-side" policy environments fixing the information environment, the expected profits of the firm remain almost constant. Although, as table 4.10 shows, the firm realizes higher profits at the star rating levels where the bonuses are awarded, it is expected that they will remain the same.

The fifth column of table 4.4 displays the levels of the expected consumer surplus that arise under the six scenarios of interest. Consumer surplus increases as consumers become more informed. This increase is explained by two reasons that are also conveyed in table 4.11. First, as consumers become more informed quality increases, and thus consumer surplus increases since consumer utility is increasing in quality. Second, as consumers become more informed, they base their choice decisions on the realized star rating levels, ignoring their prior beliefs. Hence, the firm prices its plan appropriately, leading to lower prices. Since consumer utility is decreasing in prices, the consumer surplus increases. Interestingly, the increase in consumer surplus is not very high. When profits decrease by 30% when we move from a fully uninformed, $\lambda = 0$, to a fully informed, $\lambda = 1$, environment, consumer surplus increases only by 9.68%. Overall, this increase is not enough to cover the decrease of the firm profitability, and consequently, the overall welfare decreases. This controversial result implies that the information the SRS provides to consumers cannot lead to significant welfare improvement. Comparing the levels of consumer surplus that arise under the different "supply-side" policy environments fixing the information environment, the expected consumer surplus increases slightly. This increase is explained mainly by the lower prices the firm offers when it receives the bonuses at the higher star rating levels, as shown in table 4.8.

The last column of table 4.4 shows that the government expenditures increase as consumers become more informed only in the case where the "supply-side" QBP is implemented. This is not surprising. As consumers become more informed, quality increases, and thus the bonus subsidies increase.

4.3.2 Evaluation of the current policies

Star Rating System: Table 4.5 displays how the SRS affected the MA market. Informing 22% of the Medicare population, the SRS led to a net increase in quality by 0.12. This increase raised the monthly cost of the firm per enrollee it serves by $3, while it did not affect significantly either the expected equilibrium prices or market shares. The monthly expected profits of the firm decreased by $2.5, the consumer surplus increased by $0.4, while the government expenditures remained the same. Overall, these changes imply that, although quality increased, the overall welfare decreased due to the higher decrease in the expected profits compared to the increase in the consumer surplus.

Quality Bonus Program: Table 4.6 displays how the QBP affected the MA market given that the proportion of consumers that is informed is $\lambda = 0.22$. The implementation of the QBP led to a net increase in quality of 0.19. This increase raised the monthly cost of the firm per enrollee it serves by $5, decreased the expected monthly price by $1.64, and had no effect on the market shares. The monthly expected profits of the firm only increased by \$0.31 per enrollee it serves, the consumer surplus increased by \$1.24, while the government expenditures increased by \$6. Interestingly, of the extra \$6 the government spends for the realized quality improvement, only 25% of it is justified by the realized increase of the total welfare. The remaining 75%, which amounts to an annual level of \$1.2b, is not rationalized.[12]

Discussion: A caveat of my analysis is that it does not address the credibility of the star ratings as signals of quality or whether the dimension of quality the star ratings signal on is valued by consumers. In my analysis, I have not separately identified the weight, w, consumers put on the information they receive through the star ratings, r, from their willingness to pay for quality, q. Hence, I do not know whether the small size of the consumer surplus improvement is due to the low quality of the star ratings as signals of quality or the low consumer willingness to pay for this dimension of quality. In either case, the total value consumers allocate to star ratings does not increase consumer surplus enough to improve welfare overall.

A question arises regarding the utilization of the expenditure level that is not rationalized by the realized welfare improvement. Does the QBP generate some positive externality that justifies this level of expenditure, or is this level just wasted? There are a few different forms that could reflect a potential positive externality. For example, one could show the improved health of consumers resulting from the realized quality improvement. Interestingly, [1]

[12] $1.2b = 476 \times 12$ (months) $\times 24m$ (MA population) .

provide evidence of the opposite pattern: star ratings are positively correlated with the mortality measure they use as a general proxy of health outcomes, suggesting that higher-ranked plans have higher mortality rates. No evidence regarding other health outcomes and their relation to the star ratings has been documented. Another form of positive externality could refer to the market failure the consumer misinformation creates. When consumers think that quality is higher than what it already is, they are misled. By increasing quality, the QBP brings the equilibrium quality levels closer to consumer beliefs and might correct for this failure. In either case, an extra star rating should generate an amount of $23 per month per enrollee a plan serves to justify the amount that is not rationalized by the observed welfare increase. My study remains agnostic regarding whether this level of positive externality is actually generated or whether the expenditures that are not rationalized by welfare improvement constitute a loss.

Specifications	(1)	(2)	(3)	(4)
Variable				
κ	.063	.068	.077	.087
q_0	-4.909	-4.753	.935	-8.144
γ_{drug}	1.269	1.308	1.431	1.485
γ_{vision}	.086	.086	.207	.180
γ_{dental}	.154	.111	.233	.199
γ_{hear}	-.030	.073	.103	.283
Willingness to pay (in$/month) for:				
drug covg	210.00	218.00	187.72	171.83
vision covg	14.33	14.33	27.12	20.89
dental covg	25.66	18.5	30.56	23.10
hear covg	-5.00	12.16	13.51	32.86
Mean own-price elasticity	-.857	-.921	-1.032	-1.164
Year FEs	N	Y	N	Y
Contract FEs	Y	Y	Y	Y
IVs	N	N	Y	Y
Observations	61146	61146	61146	61146

Table 4.1: Full demand dstimates

	(1)	(2)	(3)	(4)
τ_r	-.020***	.583***	.041***	.382***
	(.001)	(.020)	(.003)	(.029)
τ_{r^2}		-.082***		-.046***
		(.002)		(.004)
τ_{drug}	.229***	.223***	.211***	.211***
	(.003)	(.003)	(.004)	(.004)
τ_{vision}	.025***	.021***	.025***	.022***
	(.002)	(.002)	(.004)	(.004)
τ_{dental}	.055***	.059***	.044***	.046***
	(.002)	(.002)	(.003)	(.003)
τ_{hear}	-.007***	.005***	-.017***	-.008**
	(.002)	(.002)	(.003)	(.003)
τ_0	6.401***	5.320***	6.194***	5.583***
	(.007)	(.036)	(.012)	(.052)
Market FEs	N	N	Y	Y
Adj. R^2	.125	.140	.133	.138
Observations	61146	61146	61146	61146

Table 4.2: Marginal cost parameter estimates

	Estimate	Standard error
σ_η	1.455	(.035)
Observations	61146	

Table 4.3: Noise of star ratings

		$\tilde{q}_j^*$	$E[\pi_j]$	$E[CS_i]$	$E[K_j]$
w/o QBP	$\lambda = 0.00$	2.075	30.12	16.73	757
	$\lambda = 0.22$	2.195	27.63	17.13	757
	$\lambda = 1.00$	2.899	21.20	18.35	757
w/ QBP	$\lambda = 0.00$	2.221	30.24	17.80	762
	$\lambda = 0.22$	2.386	27.94	18.37	763
	$\lambda = 1.00$	3.250	23.92	20.57	770

Table 4.4: Market equilibrium outcomes

	$\tilde{q}_j^*$	$E[C_j]$	$E[p_j]$	$E[s_j]$	$E[\pi_j]$	$E[CS_i]$	$E[K_j]$
$\lambda = 0$	2.075	667	47.77	.218	30.12	16.73	757
$\lambda = 0.222$	2.195	670	47.83	.204	27.63	17.13	757
Net changes	.120	3	.060	-.014	-2.490	.400	0

Table 4.5: Evaluation of the SRS

	$\tilde{q}_j^*$	$E[C_j]$	$E[p_j]$	$E[s_j]$	$E[\pi_j]$	$E[CS_i]$	$E[K_j]$
w/o QBP	2.195	670	47.83	.204	27.63	17.13	757
w/ QBP	2.386	675	46.19	.206	27.94	18.37	763
Net changes	.191	5	-1.640	.002	.310	1.240	6

Table 4.6: Evaluation of the QBP

		$\tilde{q}_j^*$	$E[\pi_j]$	$E[CS_i]$	$E[K_j]$
w/o QBP	$\lambda = 0.00$	2.085	27.47	16.78	810
	$\lambda = 0.22$	2.177	26.95	17.04	810
	$\lambda = 1.00$	2.663	21.46	1.75	810
w/ QBP	$\lambda = 0.00$	2.223	27.64	17.56	814
	$\lambda = 0.22$	2.346	26.17	17.90	816
	$\lambda = 1.00$	2.969	22.80	18.97	821

Table 4.7: Average market equilibrium outcomes for monopolistic markets

		Star Ratings									
		0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
w/o QBP	$\lambda = 0.00$	11	21	31	42	53	65	77	90	103	117
	$\lambda = 0.22$	7	17	28	39	51	64	77	90	104	118
	$\lambda = 1.00$	0	5	18	32	46	61	76	91	107	123
w/ QBP	$\lambda = 0.00$	11	21	31	42	53	65	77	58	70	83
	$\lambda = 0.22$	7	17	28	39	51	64	77	58	71	85
	$\lambda = 1.00$	0	5	18	32	46	61	76	60	75	92

Table 4.8: Equilibrium monthly prices at different star rating levels

		Star Ratings									
		0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
w/o QBP	$\lambda = 0.00$	.272	.255	.238	.221	.205	.190	.175	.161	.149	.137
	$\lambda = 0.22$	.249	.235	.221	.208	.196	.184	.174	.164	.155	.147
	$\lambda = 1.00$	.159	.159	.159	.160	.162	.165	.168	.172	.176	.182
w/ QBP	$\lambda = 0.00$	.272	.255	.238	.221	.205	.190	.175	.199	.183	.168
	$\lambda = 0.22$	.249	.235	.221	.208	.196	.184	.174	.201	.190	.180
	$\lambda = 1.00$	.159	.159	.159	.160	.162	.165	.168	.210	.214	.218

Table 4.9: Equilibrium market shares at different star rating levels

		Star Ratings									
		0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
w/o QBP	$\lambda = 0.00$	39.94	36.55	33.35	30.37	27.61	25.07	22.74	20.62	18.7	16.96
	$\lambda = 0.22$	35.50	32.87	30.42	28.15	26.07	24.18	22.48	20.97	19.63	18.46
	$\lambda = 1.00$	20.29	20.26	20.32	20.48	20.74	21.11	21.60	22.21	22.94	23.78
w/ QBP	$\lambda = 0.00$	39.94	36.55	33.35	30.37	27.61	25.07	22.74	26.54	23.96	21.61
	$\lambda = 0.22$	35.50	32.87	30.42	28.15	26.07	24.18	22.48	26.97	25.10	23.42
	$\lambda = 1.00$	20.29	20.26	20.32	20.48	20.74	21.11	21.60	28.49	29.12	29.87

Table 4.10: Equilibrium monthly profits of a firm for different star ratings

		Star Ratings									
		0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
w/o QBP	$\lambda = 0.00$	14.36	15.15	15.92	16.66	17.35	17.99	18.56	19.07	19.50	19.86
	$\lambda = 0.22$	14.93	15.65	16.34	16.98	17.58	18.12	18.60	19.01	19.35	16.61
	$\lambda = 1.00$	17.04	17.46	17.83	18.15	18.41	18.61	18.74	18.81	18.80	18.73
w/ QBP	$\lambda = 0.00$	14.36	15.15	15.92	16.66	17.35	17.99	18.56	24.99	25.68	25.27
	$\lambda = 0.22$	14.93	15.65	16.34	16.98	17.58	18.12	18.60	24.90	25.44	25.88
	$\lambda = 1.00$	17.04	17.46	17.83	18.15	18.41	18.61	18.74	24.59	24.60	24.53

Table 4.11: Consumer surplus at different star ratings

	Star Ratings									
	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Cost	622	635	648	662	676	690	704	719	734	750
Payment w/o QBP	757	757	757	757	757	757	757	757	757	757
Payment w/ QBP	757	757	757	757	757	757	757	794	794	794

Table 4.12: Cost and payments for different star ratings

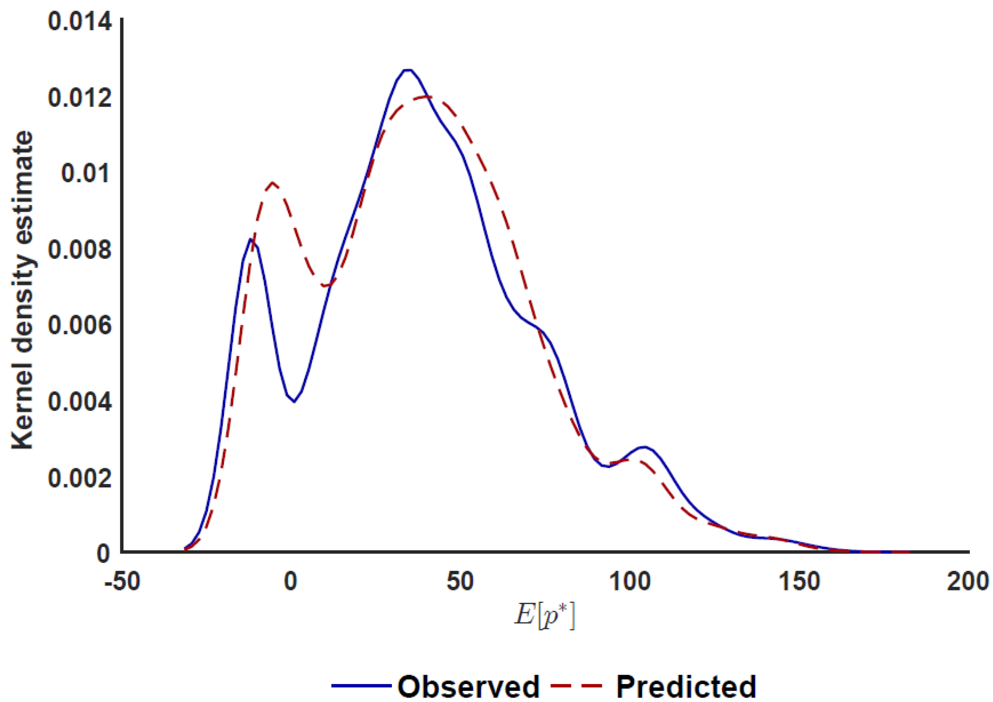


Figure 4.1: Model fit – prices

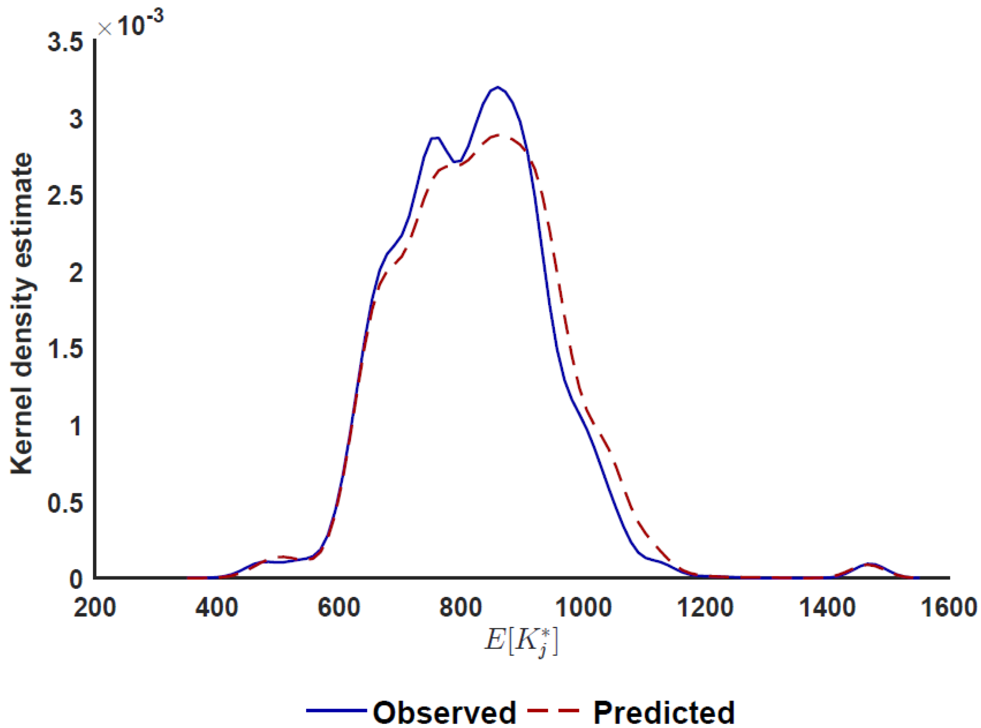


Figure 4.2: Model fit - plan payments

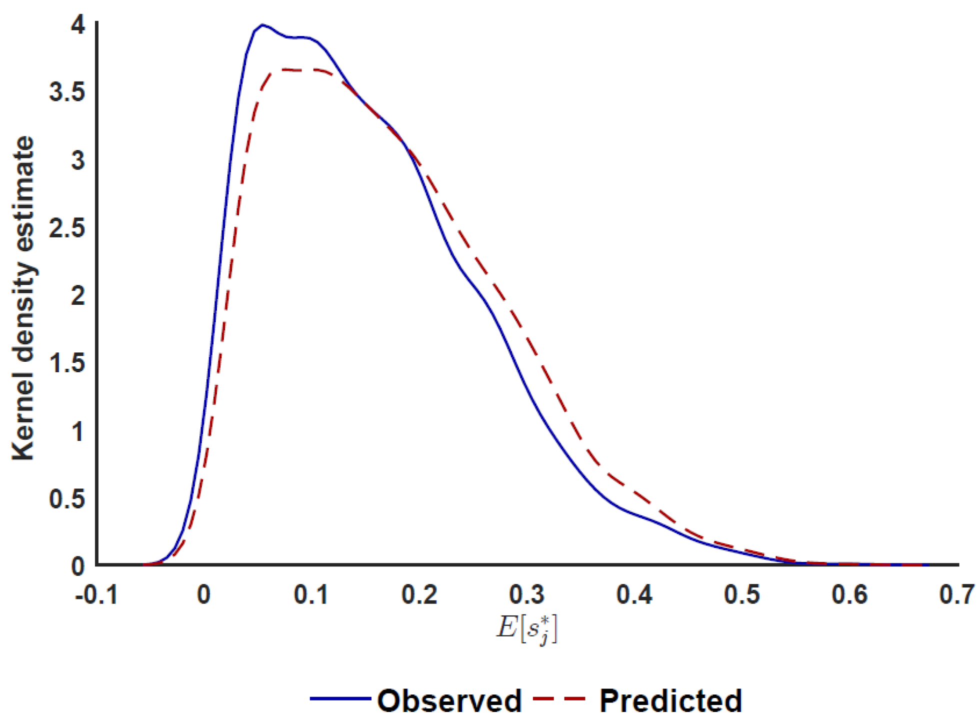


Figure 4.3: Model fit - market shares

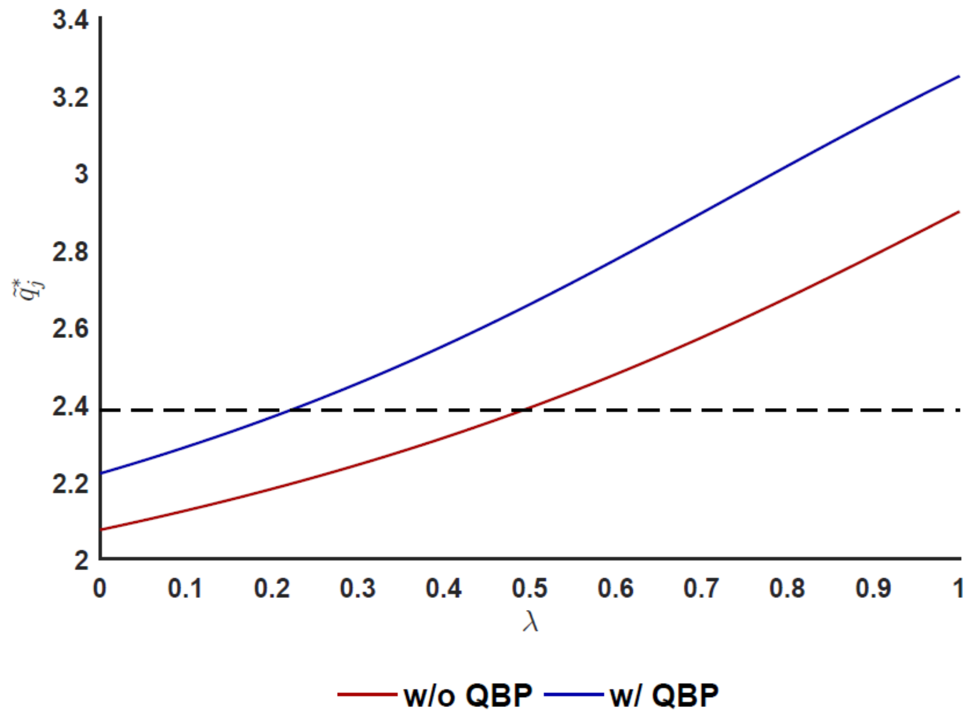


Figure 4.4: Equilibrium quality for different levels of information

Chapter 5

Conclusion

This study evaluated the relative impacts of two policies, quality disclosure, SRS, and quality bonus subsidies, QBP, on the welfare of the Medicare Advantage market. My framework combined a "demand-and supply-side" model that allowed for consumer heterogeneity with respect to preferences for and knowledge of the SRS. I built a full demand and supply equilibrium model. On the demand side, I allowed consumers to learn about quality in a Bayesian fashion after the introduction of the SRS. To separately identify consumers who were not aware of the SRS from consumers who did not care about the SRS, I surveyed 624 nationally representative Medicare beneficiaries and, I directly asked them about their knowledge with respect to the SRS. Within the survey, I conducted a conjoint analysis to elicit respondents' preferences for star ratings. I combined these stated preference with revealed preference choice data to estimate the remaining parameters of the main demand model. On the supply side, I allowed insurers to choose both prices and quality, endogenizing the information environment and the financial incentives they received.

My survey found that almost 80% of the population was *unaware* of the SRS. The estimates of the survey showed that respondents who reported they were aware of the SRS placed a monthly value of \$25 on an extra star rating, while the respondents who reported they were unaware of the SRS placed a value of \$20 per month on an extra star rating. On the supply side, I found that an extra star rating costs a firm slightly less than consumer willingness to pay for it, and the firms' choices were governed by noise, which prevented them from perfectly predicting the final level of quality that would arise. My welfare analysis showed that both the SRS and the QBP improved quality. Interestingly, if 50% of the population was aware of the SRS, the "demand-side" policy itself would be enough to lead to the levels of quality the market is currently experiencing under the implementation of both SRS and QBP. The most striking result was that 75% of the government expenditures for providing quality bonuses are not rationalized by welfare improvement.

I limit my attention to analyzing welfare as it is traditionally defined as the sum of firm profits and consumer surplus. My study remains agnostic on whether the QBP generates a potential positive externality or if it just distorts the market by wasting almost \$1.2b annually. This positive externality could take the form of a healthier Medicare population due to the higher quality the market experiences. However, [1] show that star ratings are not associated with any commonly used proxies of health outcomes in this particular market. Another potential externality of the QBP could take the form of correcting the distortions that arise when consumers have inaccurate beliefs about quality. The investigation of this possibility is left for future research.

Another simplification of my model is that it assumes that star ratings are perfectly informative. Data limitations prevent me from separately identifying how much consumers value star ratings and how much they value quality per se. Specifically, given the discrete choice model of the conjoint analysis, preferences for star ratings–captured by the coefficient β_i^s–effectively reflect a combination of consumer preferences for the actual quality–clinical outcomes and customer satisfaction–and the weight consumers put on the signals the star ratings give them. Consequently, I cannot make conclusions regarding the efficacy of star ratings as signals of quality or whether the dimension of quality the star ratings are signaling on is indeed valuable to consumers. A potential solution to address this issue would be the introduction of Bayesian assumptions in the conjoint experiment along with an additional round of surveys that would introduce variation in the prior beliefs of the survey respondents. Such a task is left for future research.

Finally, the high cost of the QBP leads to many questions that open avenues for future research. It would be interesting to investigate whether bonus schemes that depend on the relative performance evaluation of firms could be more efficient. Relative performance evaluation refers to the evaluation of an agent's performance relative to that of a peer. This type of bonus scheme can be efficient in the event that a principal, in this case the government, does not have perfect information about the agent's cost to produce quality. Introducing a scheme that assigns bonuses to firms based on their performance relative to their competitors' performance not only induces competition in quality, but also solves the problem of the asymmetric information. Evaluating this type of scheme requires a model that accounts for the interaction between the principal and the agents in the market as well as the relationship between the agents themselves.

Appendix

Appendix 1

More on the Institutional Details

1.1 More on the Star Rating System

This section describes the methodology followed by CMS for the calculation of star ratings. CMS classifies contracts into three types; MA-only that offers Part C benefits only, PDP that offers Part D benefits only, and MA-PD that offers both Part C and D benefits. Each contract type is rated over a number of performance measures (MA-only up to 32, PDP up to 15, and MA-PD up to 44[1]) that span five broad categories that are consistent with CMS's goals. These categories are the following:

1. outcomes that refer to a beneficiary's health resulting from the provided care,
2. intermediate outcomes that help move closer to true outcomes,
3. patient experience that refers to a beneficiaries' perspective on the care they received,

[1] 44 is the sum of 32 and 15 after subtracting 3 measures that overlap between MA-only and PDP contract types.

4. access that refers to any issues that may create obstacles in receiving the needed care, and

5. process that refers to the method by which health care is provided.

Every year, CMS reviews the measures constituting these categories, and depending on their reliability, potential data issues and other received feedback, it makes changes on the current measures, deletes and/or adds redundant and/or more appropriate measures, respectively.

Star ratings are reported in five different levels.[2] These are:

1. base level that reflects individual metrics comprised of numeric data (percentage scores),

2. star level that reflects star ratings calculated based on algorithms converting base level metric rates on a 5-star scale,

3. domain level that reflects a second level into which each metric grouped with similar metrics (at a star level),[3]

4. summary Part C (Part D) level that reflects metrics (at a star level) grouped together to form the Part C (Part D) summary for a contract, and

[2]This description is based on the latest version of the 2015 data available. Previous years' forms of the reported data were similar with the difference that before 2011 overall star ratings are not reported.

[3]Totally, there are 9 domains comprised of up to 47 measures. MA-only contracts are measured on 5 domains, PDP contracts on 4 domains, and MA-PD contracts on all 9 domains.

5. overall level that reflect Part C and Part D metrics (at a star level) grouped together.

The domain rate is the unweighted mean of the individual star ratings. To receive a domain rate, the contract must meet or exceed a minimum required number of individual metrics. The summary Part C, Part D, and overall MA-PD star ratings are weighted averages of the individual star ratings. To receive a Part C, and/ or a Part D summery rate, a contract must meet a minimum required number of individual metrics. For the Part C and D summary rates, half stars are also assigned to allow more variation across contracts. Lastly, for MA-PD contract types to receive an overall rate, the contract must have stars assigned to both Part C and D summary rates, and the overall star rating is calculated using a weighted average of the Part C and D summary rates. For the overall star rates, half stars are also assigned to allow more variation across contracts.

1.1.1 Recovering Continuous Levels of Quality

Over the period of 2008-2016 the algorithm changed twice; once in 2012 and then in 2016. I used the yearly algorithm as it was provided by CMS along with the individual star rating data to construct continuous levels of quality (i.e. the average summary rates). Regardless of the changes occurred during this time, generally, in the main algorithm the part C and D summary rates were calculated by taking a weighted average of the individual Part C and D metrics, respectively. CMS uses both the mean and the variance of individual metrics. Specifically, an "integration factor" was calculated and added to the mean score for the reward of high performing for a long period of time contracts. There were a few overlapping metrics between the part C and D individual metrics which were excluded for the calculation of the overall star ratings. Below, I describe the main steps of the algorithm I followed.

1. Classify contracts as "MA only", "PDP", and "MA-PD" if offering only Part C, only Part D, and both Part C and Part D benefits, respectively.

2. Generate the weights each individual metric receives.

3. Calculate the "MA-only" rate as follows:

   (a) Calculate the Part C weighted average with and without the improvement metric.

   (b) Calculate the Part C weighted variance with and without the improvement metric.

   (c) Categorize both versions of weighted average as (i) below 65th pctile, or (ii) above 65th and below 85th pctile, or (iii) above 85th pctile.

   (d) Categorize both versions of weighted variance as (i) below 30th pctile, or (ii) above 30th and below 70th pctile, or (iii) above 70th pctile.

   (e) Develop the "integration" factor for both versions (with and without improvement) depending on the mean and variance categories generated above.

(f) Add "integration" factor to the mean score for both versions.

(g) If a contract has Part C weighted average without the improvement metric less or equal to 2, keep that level of quality. If a contract has Part C weighted average without the improvement metric greater or equal to 4, keep the maximum of the Part C weighted average with and without the improvement metric. For all other cases, keep the Part C weighted average with the improvement metric.

4. Calculate the "PDP" rates following the same steps as in "MA-only".

5. Calculate the overall "MA PD" rates: follow the same steps as in "MA only" and "PDP" cases excluding the overlapping measures.

6. Create Part C, D and overall scores rounding to the nearest half star.

During the period of interest CMS changed the algorithm it used to construct the final star ratings a few times. This fact can generate concerns on whether the comparison between the distributions over time is credible. To address this issue, I also constructed a measure of continuous quality using the algorithm CMS used in 2016 and holding it constant for all the years of my analysis. Figure 1.1 displays how the distributions of this version of quality evolved over time. Overall, we observe a similar shift of the distributions to the right and a multimodal shape beginning of 2012 and becoming more distinct as we get close to 2016. It is interesting to notice the main difference between this figure and figure 2.4.The distributions in this figure are almost uni-modal closer to 2008 and become multi-model closer to 2016, whereas the distributions in figure 2.4 present a more multimodal pattern across all years.[4] The pattern of the distributions in figure 1.1 implicitly suggests that in 2008, when insurers did not know the algorithm CMS would use in 2016, they could not target overall quality very well, but getting closer to 2016 they could target better.

1.2 Identification challenge

To better understand the identification challenge, assume that the star ratings are perfectly informative, i.e. $w = 1$ and that preferences for quality, q_{kmt} are represented by a simple random coefficient, $\beta_i = \overline{\beta} \cdot \Phi_i$, where Φ_i is a discrete random variable that indicates whether consumer i cares about quality, q_{kmt} and follows the Bernoulli distribution, $\Phi_i \sim Bern(\phi)$ as follows:

$$\Phi_i = \begin{cases} 1 & \text{w/ prob. } \phi \\ 0 & \text{w/ prob. } 1-\phi \end{cases}. \tag{1.1}$$

Supposing that information given by the SRS reaches out to everyone (as it is commonly assumed in the literature), i.e. that $\lambda_i = 1 \; \forall \; i$, the utility

[4]Remember that continuous quality in figure 2.4 is adjusted to the changes CMS makes annually on the algorithm it uses.

specification takes the following form:

$$E[u_{ikjmt}|\overline{x}_{kjmt}, r_{kmt}, \theta_i, \epsilon_{ikjmt}] =$$
$$= \begin{cases} q_0 + \alpha_i p_{kjmt} + \overline{\beta}(r_{kmt} - q_{TM}) + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \Phi_i = 1 \\ q_0 + \alpha_i p_{kjmt} + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \Phi_i = 0 \end{cases}. \tag{1.2}$$

Since star ratings, r_{kmt} are perfectly informative on quality, q_{kmt}, if the consumer cares about quality, her preferences will be represented by $\overline{\beta}$, otherwise her preferences will be represented by a zero mass point as Part (A) of figure 1.2 also illustrates. However, if we take into account the fact that there are consumers that are not aware of the SRS and in case $\mu_0 = q_{TM}$, the utility specification takes the following form:

$$E[u_{ikjmt}|\overline{x}_{kjmt}, r_{kmt}, \theta_i, \epsilon_{ikjmt}, \lambda_i] =$$
$$= \begin{cases} q_0 + \alpha_i p_{kjmt} + \overline{\beta}(r_{kmt} - q_{TM}) + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \Lambda_i = 1, \Phi_i = 1 \\ q_0 + \alpha_i p_{kjmt} + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \Lambda_i = 1, \Phi_i = 0 \\ q_0 + \alpha_i p_{kjmt} + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \Lambda_i = 0, \Phi_i = 1 \\ q_0 + \alpha_i p_{kjmt} + \gamma x_{kjmt} + \xi_{kjmt} + \epsilon_{ikjmt} & \text{if } \Lambda_i = 0, \Phi_i = 0 \end{cases}. \tag{1.3}$$

Equation 1.3 clearly indicates that behind the zero mass point there are hidden more consumer types than just simply consumers who do not care about quality as Part (B) of figure 1.2 also illustrates. More importantly, since different consumer types can have different preferences for quality, not taking into account their existence can lead to biased welfare analyses. In a more flexible model where preferences for quality, q_{kmt} allow for more flexible substitution patterns it is even more cumbersome to separately identify different consumer types.

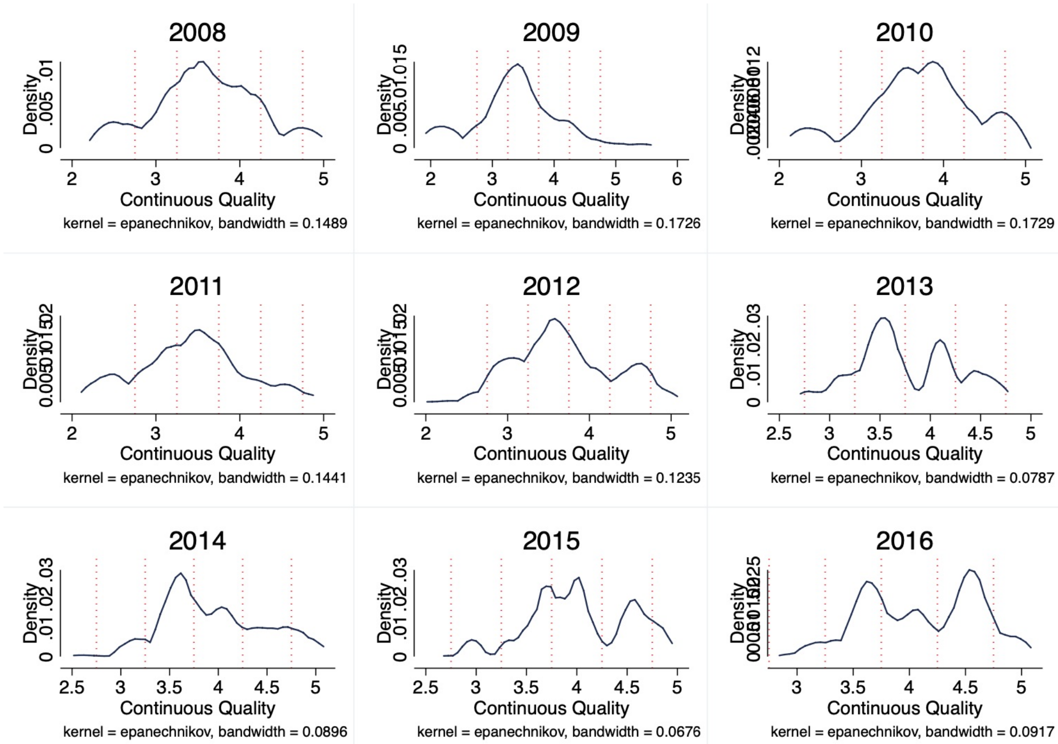
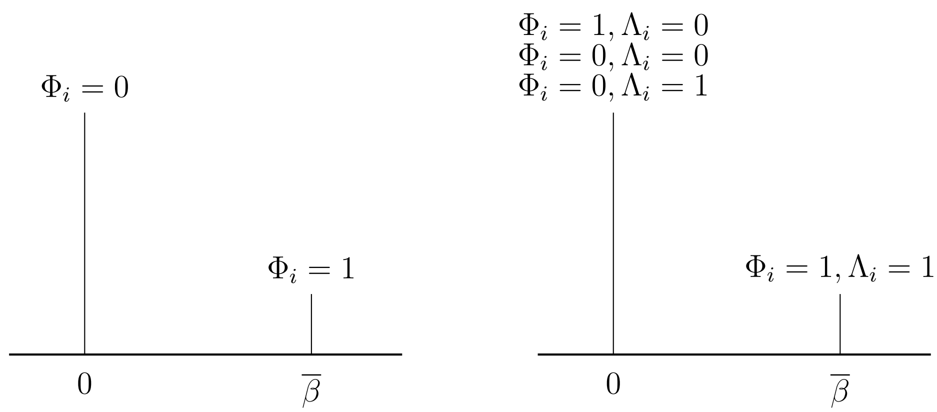


Figure 1.1: Constructed average summary rates - base year: 2016



(a) ***Naïve version*** (b) ***Flexible version***

Figure 1.2: Identification challenge

Appendix 2

More on the survey

2.1 Data quality

Respondents were guaranteed that any of the information they would provide would be kept confidential and that their responses would only be published at an aggregate level. With the exception of the demographic questions for which people tend to be sensitive, all respondents were forced to answer all questions. Although this increased the risk of respondent dropping out of the survey, it also guaranteed that the collection of the information would be consistent and the data set fairly balanced.

To test the quality of my survey data I followed a series of tests. First, I checked that indeed all respondents matched the screening questions. Second, I checked whether the state of their current residence was in alliance with the zip-code they provided. Third, I excluded respondents who spent less than 1/3 of the average length the survey would take a respondent. Fourth, for a group of respondents that were at the low end of the distribution of the length they spent to take the survey, I checked the extent to which they were choosing consistently the same options in the conjoint experiments and/or whether consistently chose the same level of importance in all the plan characteristics of the relevant question I provided. Lastly, I checked whether the answers they provided in the open-ended questions were indeed related to the topic and did not reveal any confusions.

Concerns regarding potential inattention with respect to the survey might arise, overall. To minimize such concerns, I set the option to advance to the next question not to appear for the first few seconds each question is available and I recorded how long it took each respondent to take the entire survey and to respond each separate question, as well. The average length of the survey was 8.4min. The average time respondents spent on the first question that explain them the context and the goal of the survey was 40sec. The average time they spent to answer the awareness question was 30sec. Lastly, the average time they spent on the conjoint experiment was 1-2min with most of the time spent on the first two questions.

2.2 Survey Questionnaire

1. What is your age?
2. In which state do you currently reside?
3. *See Figure 3.2*
4. Sometimes Medicare-eligible individuals get their health insurance plans partially or fully paid for by their current or previous employer. Please read the options below carefully, and choose the one that applies to you.

- My former/current employer pays for/helps me pay for my health insurance plan.
- My former/current employer does not help me pay for my health insurance plan

5. Have you ever been enrolled in a Medicare Advantage plan?
   - Yes
   - No
   - I do not recall
6. Please indicate when you first enrolled in Medicare Advantage.
   - 2019
   - 2018
   - 2017
   - 2016
   - 2015
   - Other. Please indicate.
   - I do not recall
7. When did you join your current Medicare Advantage plan?
   - In the last open enrollment period for 2020.
   - 1 year ago

- 2 years ago
- 3 years ago
- 4 years ago
- 5 years ago
- More than 5 years ago. Please indicate how many years ago.
- I do not recall

8. What plan were you enrolled in before?
   - A Medicare Advantage plan different from my current plan.
   - I was in Original Medicare
   - I was not eligible for Medicare
   - Other. Please explain

9. To help beneficiaries find the insurance plan that best matches their needs, Medicare rates Medicare Advantage plans on a "star" scale from 1 to 5, with higher stars indicating higher quality.
   Every year before the enrollment period begins, each plan is assigned an Overall Star Rating that indicates different levels of quality in terms of health outcomes of the people who enroll in the plan, the way plans help enrollees manage their chronic conditions, members' experiences with the plan, access to medical care, as well as customer service.

Do you remember seeing/ hearing/ reading about Overall Star Ratings for Medicare Advantage plans?

- Yes
- No

10. Where exactly did you see/hear/read about them?

- Medicare.gov
- Insurer's website/handbook
- Plan advertising/mailers
- "MedicareAndYou" handbook
- I do not recall
- Other. Please explain

11. Please indicate the Overall Star Rating for your current Medicare Advantage plan.

- 1
- 1.5
- 2
- 2.5
- 3
- 3.5

- 4
- 4.5
- 5
- I do not recall

12. You indicated that you currently know about the Star Rating System. Did you know anything about it when you originally decided to receive your Medicare coverage from Medicare Advantage?

    - Yes
    - No

13. You indicated that you currently know about the Star Rating System. Did you know anything about it when you originally decided to join your current health insurance plan?

    - Yes
    - No

14. Where exactly did you see/hear/read about the Star Rating System?

    - Medicare.gov
    - Insurer's website/handbook
    - Plan advertising/mailers
    - "MedicareAndYou" handbook

- I do not recall
- Other. Please explain

15. If Original Medicare was suddenly unavailable, and you had to choose from the existing Medicare Advantage plans in your area, would you take the Overall Star Ratings into consideration?

    - Yes
    - No

16. If your health insurance plan suddenly stopped being offered, and you had to choose from a set of other Medicare Advantage plans, would you take the Overall Star Ratings into consideration?

    - Yes
    - No

17. Why not?

    - I already know the information the Overall Star Ratings provide.
    - I do not understand the information the Overall Star Ratings provide.
    - I am not interested in the information the Overall Star Ratings give me.
    - I do not trust this Star Rating System.

- I do not find the Overall Star Ratings very informative
- Other. Please explain

18. Approximately, how much is the monthly premium of your current Medicare Advantage plan?

    - The approximate monthly premium of my current Medicare Advantage plan is: *(Please enter your answer in the box below)*
    - I do not recall

19. *See figure 3.4*

20. If the two plans presented below were identical in every other way except for the following two characteristics, which one would you choose to enroll in?

    Note: There is no right or wrong answer. You should select the option that best reflects your personal preferences.

    - **Plan1**: Monthly premium: $21 Overall Star Rating: 2
    - **Plan2**: Monthly premium: $29 Overall Star Rating: 2.5

21. If the two plans presented below were identical in every other way except for the following two characteristics, which one would you choose to enroll in?

    Note: There is no right or wrong answer. You should select the option that best reflects your personal preferences.

- **Plan1**: Monthly premium: $52 Overall Star Rating: 5
- **Plan2**: Monthly premium: $32 Overall Star Rating: 4

22. If the two plans presented below were identical in every other way except for the following two characteristics, which one would you choose to enroll in?

    Note: There is no right or wrong answer. You should select the option that best reflects your personal preferences.

    - **Plan1**: Monthly premium: $52 Overall Star Rating: 3.5
    - **Plan2**: Monthly premium: $87 Overall Star Rating: 5

23. If the two plans presented below were identical in every other way except for the following two characteristics, which one would you choose to enroll in?

    Note: There is no right or wrong answer. You should select the option that best reflects your personal preferences.

    - **Plan1**: Monthly premium: $0 Overall Star Rating: 3
    - **Plan2**: Monthly premium: $49 Overall Star Rating: 5

24. Did someone help you make your annual health insurance choice for 2020?

    - Yes
    - No

25. Who helped you?

    - A counselor from the State Health Insurance Assistance Program (SHIP)
    - Medicare helpline
    - Family member
    - Independent agent/broker
    - My insurance company
    - Other. Please explain

26. How did they help you?

    - They suggested/chose a plan for me
    - They showed me the Medicare Plan Finder website at Medicare.gov/find-a-plan
    - They helped me use the Medicare Plan Finder website at Medicare.gov/find-a-plan.
    - Other. Please explain

27. *See Figure 3.9*

28. Is there anything else you that you took into consideration before making your annual health insurance choice for 2020?

    - Yes. Please specify

- No

29. On a scale from 1 to 100, how satisfied are you with your current Medicare Advantage plan?

30. On a scale from 1 to 100, how satisfied are you with the health care you receive from Original Medicare?

31. How would you describe your current health?

    - Excellent
    - Very good
    - Good
    - Fair
    - Poor

32. What is your gender?

    - Male
    - Female
    - Other

33. What is your ethnicity?

    - White
    - Hispanic or Latino/a

- Black or African American
- Native American or American Indian
- Asian/Pacific Islander
- Other. Please specify

34. What is your marital status?
    - Married
    - Widowed
    - Divorced
    - Separated
    - Never married

35. How much do you spend on average every month on housing, utility bills, food, transportation, healthcare, and other common leisure activities?
    - Less than $1800
    - $1801 - $2500
    - $2501 - $3400
    - $3401 - $4200
    - $4201 - $5000
    - Above $5000

36. What is the highest level of education that you have?

- No schooling
- 8th grade/less
- 9 - 11 grades
- High school
- Bachelor's degree
- Graduate degree
- Other. Please specify

37. What is the 5-digit ZIP code of your current residence?

38. Is there anything else at all that you would like to share about Medicare Advantage?

39. Is there anything else at all that you would like to share about Original Medicare?

2.3 Answers to open-ended questions

Question: Is there anything else at all that you would like to share about Medicare Advantage?

- I like the convenience of having only one plan to keep up with
- I like my plan
- The advantages far outweigh the extra cost

- Excellent program
- Some questions are difficult to get answers from the plan quickly. This can be extremely confusing when enrolling. At my age I feel lucky to understand how this process works but feel advocates would be needed for older folks
- Our income is very low so we must choose plans with zero monthly premiums. However, the copays have gotten increasingly more expensive and more frequent for us as our chronic conditions worsen with age and we are now at the point of picking and choosing which specialist visits we must schedule and which are "optional" (none, really) and have to be skipped due to the cost.
- I am glad that there are several local Medicare Advantage plans to choose from that I do not have to pay anything above the monthly premium that is taken out of my Social Security check
- Need more plans for Hawaii county
- Couldn't believe I could only receive benefits in Palm Beach County
- Never have used the service
- I am extremely happy with the provider I have chosen
- I would like to be able to talk with someone who is NOT attempting to sell me on a plan

- Don’t get “X” plan!
- I do not want Medicare totally privatized
- I get excellent customer service and I am grateful for that
- Terrific
- Very good for the cost. Mine is $0
- Great plan
- That it has been a good plan for me
- It also has been very helpful
- Pay less premium monthly
- It should not cost so much and the deductible should not be so high
- Nothing I can think of right off the top of my head
- I love it, I don’t pay a monthly premium in my area
- It’s worth the small cost for the premium coverage
- I had two heart attacks last year and had it not been for my insurance, my bill would have been over $150000
- The plans with very low premiums are not worth the money
- Excellent coverage

- Congress better not mess with it
- Do not yet have enough trust in them to use
- Would like it to cover dental
- I still do not understand stars
- Yes, great coverage and picks up what Medicare doesn't cover
- What good is it if it costs more than its worth?
- Medicare advantage is much better than just Medicare
- I do not understand why it is not marketed more aggressively
- Show offer eye and dental
- Every thing has a co-pay now
- I like my plan
- I am happy with it
- Overall it is a better alternative to "Original Medicare"
- Not at all impressed especially with doctor choices unless cost is of the utmost importance to purchaser
- No except they might try bargaining with the pharmaceutical industry so that people could afford to take the medications they require

- The little I know about it is that it too complex and said to be ever-changing
- Like plan
- Hospitals and medical reps. should accept any paying insurance
- I am very satisfied with my plan
- I have a great plan
- My only problem has been that the network is very restrictive
- The main variable to me is how much I trust the company providing it and how fair they are with their coverages. I have had "X" for 11 years and they are very easy to work with, easy to talk to, fair and balanced
- Medicare advantage works great for me
- I am very pleased with my Medicare Advantage plan
- I am happy with it
- Wonderful plan
- Has worked very well for me
- Easier to use than Original Medicare
- I feel they need to offer much better coverage for vision & dental
- You will need it because Medicare coverage is bare bones

- I like my plan
- It varies by county - you cannot pin hole the advantage programs into a lump category
- The plan that I have is great
- Thank you for the opportunity to be able top share my feedback for this survey
- Great plan
- “X” plan is best I’ve found
- My plan has been great it pays for a lot of the routine items but I bought a policy with extra benefits sadly not everyone can afford that type of plan
- I appreciate Medicare Advantage - hope it lasts
- I like my $131.00 monthly reduction
- It is health insurance
- We began with “X” plan, and it is now “Y” plan
- I would like more in-network providers.
- Great option if money is an issue
- It is a must to have

- Dental coverage is restricted to a HMO
- Does it live up to the hype
- I love it
- I think that it is a good idea and affordable
- I can't think of anything else
- I like what I have

Question: Is there anything else at all that you would like to share about Original Medicare?

- It's not free
- It is a great program
- Excellent program
- I've never had Original Medicare
- Does not cover dental nor vision and only minimal other medical
- I am glad that there are other options available that provide more benefits that what Original Medicare offers
- No - other than the cost & yearly increase
- gGad it is here for me

- The “gap” is a nuisance
- It never paid for any of my care
- Original Medicare is so much better than “X” Advantage Plan. I got much better coverage with the Original
- I love my plan with “X”.
- I am concerned because I am withdrawing my IRA and now my Medicare insurance is almost doubled
- Doesn’t cover enough. Should cover custom shoe orthotics even if you don’t have diabetes
- Should also make payments on my health care. I pay a premium for this
- Definitely not as good as Medicare Advantage
- It has been very helpful
- No, other than I am very happy with what I have
- It should not cost so much
- Requires extremely expensive MidiGap policy like what I used to have over 6 years ago
- Not as complete as possible
- It is ok

- Hope it stays for retirees only. Should not be part of national health care plan
- It works fine for me
- It is terrible
- I really would like to see a care system that is regulated across the board which would weed out competition and high cost
- I think it is a good plan except for giving coverage to those who do not earn it
- Cost of premiums raises more than SS increases. Hard to keep up
- Wish they could bargain with the pharmaceutical companies to reduce prescription prices.
- Original Medicare with good supplemental insurance has resulted in mp co-pays and less than $500 out of pocket total in the 19 years I have had this type of insurance. I would rather budget for have total insurance when I need it than have to deal with co-pays and bills afterward
- It's an excellent program, I am amazed at how well it covers everything
- I wish there was more included with Medicare benefits
- My health care plan is excellent
- Medicare was good Medicare Advantage is better

- Would seem to be more expensive and more complicated than my Medicare Adv. plan
- It does not cover some very important procedures nor does it cover dental procedures
- Thank you for the opportunity to be able to provide my feedback for this survey
- No thank you
- Medicare is great but not all doctors want on to the plan which is not a fair system to the consumers it should be treated like any other insurance plan
- I need to understand more about the 80% coverage
- I do have a supplemental
- Happy
- Very complicated
- Wish drug plan was included. Plans available are not very good

Question: Is there anything else that you took into consideration before making your annual health insurance choice for 2020?

- Portability

- Out of network fees and acceptance
- Company increased benefits
- Annual deductible and my primary care physician
- I selected a different "X" plan because it offered an over the counter benefit that the previous "Y" plan did not offer
- Yes, drug costs are big for me one year I selected a carrier that exempted my insulin from the donut hole I just paid regular monthly co pay instead of the rip off price
- I did not have a choice. No other company will accept me because of current medical diagnosis
- Doctors who are in network and their ratings
- I have chosen to stay with the same supplemental insurance company
- I did not want to change
- Are my doctors in the network
- Main concern was network of providers
- Customer service
- Only that all my doctors participate in the plan
- Coverage outside US for emergency

- Just drug coverage
- Formulary
- Agent advice
- Past experience with the company
- Current doctor is in the plan
- Hearing aid coverage
- Location of doctors and facilities and hospitals
- Doctors and hospitals included in the plan
- Nothing that I have not considered
- Drugs
- I have been satisfied with the coverage and price along with the ease of accessing MDs
- Hospitals in network were even more important to me than doctors
- Continuity of coverage
- My preferred md in plan
- My overall satisfaction with the plan over the prior several years
- Areas of coverage

- Overall cost for me, inclusion of doctors I want
- Prevention benefits
- Dental
- That my current doctors are in the plan
- Transportation provided
- What a pain it was to review all available options and insurers
- Making sure that providers that I use are covered by the plan
- How close was a doctor to me
- Had them the year prior and was very satisfied
- Has local doctors that I use and no copay for primary doctor
- Yes my doctor was on this plan
- I wanted to keep my agent who is extremely capable knowledgeable and approachable
- The provider network
- Was my doctor in the plan
- Having my choice of health care providers Also if my coverage is available in other areas of the country if I choose to travel

- "X" Clinic accepts our insurance
- Doctors who accept medicare
- Choice of doctors and hospitals wherever I may travel in this country
- Ease of use and acceptance by my providers
- Recommendation from friends and family
- The doctor clinic that I go to accept my insurance with no copay
- PPO offered and enrolled in
- No out of pocket monthly premiums
- Location of providers
- The network of available doctors
- Location of facilities and physicians
- I made my choice not because it offered the best plan but because this is what I could afford
- Vision and gym membership
- Quality care is not worth a small savings to me I can afford the better plan
- Sizable monthly reduction

- Location of service
- My doctor has to be in the network
- Wanted a plan with no deductibles no copays no referrals needed
- Rx covered
- Coverage and copay for hospital inpatient and skilled nursing facility
- Good Doctors
- The convenience of continuity
- Cost and benefits offered
- if my doctor was in network
- Complete coverage but expensive
- Availability and location of service

Bibliography

[1] J. Abaluck, M.C. Bravo, P. Hull, and A. Starc. Mortality effects and choice across private health insurance plans. *Working paper*, July 2020.

[2] J. Abaluck and G. Compiani. A method to estimate discrete models that is robust to consumer search. *Working paper*, March 2020.

[3] J. Abaluck and J. Gruber. Choice inconsistencies among the elderly: Evidence from plan choice in the medicare part d program. *American Economic Review*, 101(4):1180–1210, June 2011.

[4] J. Abaluck and J. Gruber. Evolving choice inconsistencies in choice of prescription drug insurance. *American Economic Review*, 106(8):2145–2184, August 2016.

[5] D. A. Ackerberg. Advertising, learning, and consumer choice in experience good markets: An empirical examination. *International Economic Review*, 44(3):1007–1040, August 2003.

[6] N. Aizawa and Y.S. Kim. Advertising and risk selection in health insurance markets. *The American Economic Review*, 108(3):828–867, March 2018.

[7] G. A. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 85(3):488–500, August 1970.

[8] K.J. Arrow. Uncertainty and the welfare economics of medical care. *The American Economic Review*, 53(5):941–973, December 1963.

[9] N.D. Beaulieu. Quality information and consumer health plan choices. *Journal of Health Economics*, 21:43–63, 2002.

[10] M. Ben-Akiva, M. Bradley, T. Morikawa, J. Benjamin, T. Novak, H. Oppewal, and V. Rao. Combining revealed and stated preferences data. *Marketing Letters*, 5(4):335–349, October 1994.

[11] M. Ben-Akiva, D. McFadden, and K. Train. Foundations of stated preference elicitation: Consumer behavior and choice-based conjoint analysis. *Foundations and Trends in Econometrics*, 10(1-2):1–144, 2019.

[12] S. Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2):242–262, Summer 1994.

[13] S. Berry and P. Jia. Tracing the woes: An empirical analysis of the airline industry. *American Economic Journal: Microeconomics*, 2(3):1–43, August 2010.

[14] S. Berry, J. Levinshon, and A. Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, July 1995.

[15] P. Bordalo, N. Gennaioli, and A. Shleifer. Salience theory of choice under risk. *The Quarterly Journal of Economics*, 127(3):1243–1285, August 2012.

[16] P. Bordalo, N. Gennaioli, and A. Shleifer. Salience and consumer choice. *Journal of Political Economy*, 121(5):803–843, October 2013.

[17] Zach Y. Brown. An empirical model of price transparency and markups in health care. *Working paper*, 2018.

[18] Z.Y. Brown and J. Jeon. Endogenous information acquisition and insurance choice. *Working paper*, November 2019.

[19] D. Brownstone, D.S. Bunch, and K. Train. Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. *Transportation Research*, Part B(34):315–338, 2000.

[20] M. Cabral, M. Geruso, and N. Mahoney. Do larger health insurance subsidies benefit patients or producers? evidence from medicare advantage. *American Economic Review*, 108(8):2048–2087, August 2018.

[21] A. Caplin and M. Dean. Revealed preference, rational inattention, and costly information acquisition. *The American Economic Review*, 105(7):2183–2203, July 2015.

[22] F. Carlsson. Design of stated preference surveys: Is there more to learn from behavioral economics. *Environ Resource Econ*, 46:167–177, 2010.

[23] M. Catalano, B. Lo Casto, and M. Migliore. Car sharing demand estimation and urban transport demand modelling using stated preference techniques. *European Transport*, (40):33–50, 2008.

[24] E. Cherchi and D.A. Hensher. Workshop synhesis: Stated preference surveys and experimental design, an audit of the journey so far and future research perspectives. *Transportation Research Propedia*, 11:154–164, 2015.

[25] M. Chernew, G. Gowrisankaran, C. McLaughlin, and T. Gibson. Quality and employers' choice of health plans. *Journal of Health Economics*, 23(3):471–492, May 2004.

[26] M. Chernew, G. Gowrisankaran, and D.P. Scanlon. Learning and the value of information: Evidence from health plan report cards. *Journal of Econometrics*, 144(1):156–174, May 2008.

[27] A.T. Ching. A dynamic oligopoly structural model for the prescription drug market after patent expiration. *International Economic Review*, 51(4):1175–1207, November 2010.

[28] G.S. Crawford and M. Shum. Uncertainty and learning in pharmaceutical demand. *Econometrica*, 73(4):1137–1173, July 2005.

[29] V. Curto, L. Einav, J. Levin, and J. Bhattacharya. Can health insurance competition work? evidence from medicare advantage. *Journal of Political Economy (R and R)*, 2018.

[30] L. Dafny and D. Dranove. Do report cards tell consumers anything they don't already know? the case of medicare hmos. *RAND Journal of Economics*, 39(3):790–821, Autumn 2008.

[31] C.M. Dalton, G. Gowrisankaran, and R. Town. Salience, myopia, and complex dynamic incentives: Evidence from medicare part d. *Technical Report, National Bureau of Economic Research*, 2018.

[32] M. Darden and I.M. McCarthy. The star treatment: Estimating the impact of star ratings on medicare advantage enrollments. *Journal of Human Resources*, 50(4):980–1008, Fall 2015.

[33] A. Deaton. Price elasticities from survey data: Extensions and indonesian results. *Journal of Econometrics*, 44(3):281–309, 1990.

[34] F. Decarolis and A. Guglielmo. Insurer's response to selection risk: Evidence from medicare enrollment reforms. *Journal of Health Economics*, 56:383–396, 2017.

[35] F. Decarolis, A. Guglielmo, and C. Luscombe. Open enrollment periods and plan choices. *Working paper*, 2017.

[36] F. Decarolis, M. Polyakova, and S. Ryan. Subsidy design in privately-provided social insurance: Lessons from medicare part d. *Journal of Political Economy (forthcoming)*, February 2019.

[37] M.J. Dickstein. Efficient provision of experience goods: Evidence from antidepressant choice. *Working paper*, 2018.

[38] M. Duggan, A. Starc, and B. Vabson. Who benefits when the government pays more? pass-through in the medicare advantage program. *Journal of Public Economics*, 141:50–67, 2016.

[39] Abe Dunn. The value of coverage in the medicare advantage insurance market. *Journal of Health Economics*, 29(6):839–855, December 2010.

[40] D. Earnhart. Combining revealed and stated data to examine housing decisions using discrete choice analysis. *Journal of Urban Economics*, 51(1):143–169, January 2002.

[41] T. Erdem and M.P. Keane. Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science*, 15(1):1–20, 1996.

[42] T. Erden, M.P. Keane, and B. Sun. A dynamic model of brand choice when price and advertising signal product quality. *Marketing Science*, 27(6):1111–1125, November 2008.

[43] K.M.M. Ericson. Consumer inertia and firm pricing in the medicare part d prescription drug insurance exchange. *American Economic Journal: Economic Policy*, 6(1):38–64, February 2014.

[44] Y. Fan. Ownership consolidation and product characteristics: A study of the us daily newspaper market. *The American Economic Review*, 103(5):1598–1628, August 2013.

[45] M. Fioretti and H. Wang. Quality, quality payments, and risk selection in private medicare. *Working paper*, 2018.

[46] A. Goolsbee and A. Petrin. The consumer gains from direct broadcast satellite and the competition with cable tv. *Econometrica*, 72(2):351–281, March 2004.

[47] P.E. Green, A.M. Krieger, and Y. Wind. Thirty years of conjoint analysis: Reflections and prospects. *Interfaces (Marketing Engineering)*, 31(3):56–73, 2001.

[48] M. Grennan and R. Town. Regulating innovation with uncertain quality: Information, risk, and access in medical devices. *NBER Working Paper*, 2019.

[49] B. Handel, J.T. Kolstad, and J. Spinnewijn. Information frictions and adverse selection: Policy interventions in health insurance markets. *Review of Economics and Statistics, forthcoming*, 2018.

[50] B. Handel and J. Schwartzstein. Frictions or mental gaps: What's behind the information we (don't) use and when do we care? *Journal of Economic Perspectives*, 32(1):155–178, Winter 2018.

[51] B.R. Handel. Adverse selection and inertia in health insurance markets: When nudging hurts. *The American Economic Review*, 103(7):2643–2682, December 2013.

[52] B.R. Handel and J.T. Kolstad. Health insurance for “humans”: Information frictions, plan choice, and consumer welfare. *The American Economic Review*, 105(8):2449–2500, August 2015.

[53] F. Heiss, D. Mcfadden, and J. Winter. *Mind the Gap! Consumer Perceptions and Choices of Medicare Part D Prescription Drug Plans*, chapter Reseearch Findings in the Economics of Aging, pages 413–481. The University of Chicago Press.

[54] K. Ho, J. Hogan, and F.S. Morton. The impact of consumer inattention on insurer pricing in the medicare part d program. *RAND Journal of Economics*, 48(4):877–905, Winter 2017.

[55] G.Z. Jin and P. Leslie. The effect of information on product quality: Evidence from restaurant hygiene grade cards. *The Quarterly Journal of Economics*, 118(2):409–451, May 2003.

[56] G.Z. Jin and A.T. Sorensen. Information and consumer choice: The value of publicized health plan ratings. *Journal of Health Economics*, 25(2):248–275, March 2006.

[57] A.C. Johnston. Teacher utility, separating equilibria, and compensation structure: Evidence from a discrete-choice experiment. *Working paper*, March 2019.

[58] J.D. Ketcham, C. Lucarelli, E.J. Miravete, and M.C. Roebuck. Sinking, swimming, or learning to swim in medicare part d. *American Economic*

Review, 102(6):2639–2673, October 2012.

[59] J.D. Ketcham, C. Lucarelli, and C.A. Powers. Paying attention or paying too much in medicare part d. *American Economic Review*, 105(1):204–233, January 2015.

[60] J.T. Kolstad and M.E. Chernew. Quality and consumer decision making in the market for health insurance and health care services. *Medical Care Research and Review*, 66(1):28S–52S, February 2009.

[61] E.P. Kroes and R.J. Sheldon. Stated preference methods: An introduction. *Journal of Transport Economics and Policy*, 22(1):11–25, January 1998.

[62] T.J. Layton and A.M. Ryan. Higher incentive payments in medicare advantage's pay-for-performance program did not improve quality but did increase plan offerings. *Health Services Research*, 50(6), December 2015.

[63] G. Lewis and G. Zervas. The welfare impact of consumer reviews: A case study of the hotel industry. *Working paper*, 2016.

[64] J. Lustig. Measuring welfare losses from adverse selection and imperfect competition in privatized medicare. *Manuscript*, 2010.

[65] T. Madeira. The cost of removing deadlines: Evidence from medicare part d. *JMP*, January 2015.

[66] J. J. McCall. Economics of information and job search. *The Quarterly Journal of Economics*, 84(1):113–126, February 1970.

[67] I.M. McCarthy and M. Darden. Supply-side responses to public quality ratings. *American Journal of Health Economics*, 3(2):140–164, Spring 2017.

[68] J. McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, 2008.

[69] D. McFadden. *Handbook of Choice Modelling*, chapter The New Science of Pleasure: Consumer Choice Behavior and the Measurement of Well-Being. Edward Elgar, 2013.

[70] D. McFadden, A.C. Bemmaor, F.G. Caro, J. Dominitz, B.-H. Jun, A. Lewbel, R.L. Matzkin, F. Molinary, N. Schwartz, R.J. Willis, and J.K. Winter. Statistical analysis of choice experiments and surveys. *Marketing Letters*, 16(3/4):189–196, 2005.

[71] T.G. McGuire, J.P. Newhouse, and A.D. Sinaiko. An economic history of medicare part c. *The Milbank Quarterly*, 89(2):289–332, June 2011.

[72] K. Miller. Do private medicare firms have lower costs? *mimeo*, 2015.

[73] K.S. Miller, A. Petrin, R. Town, and M. Chernew. Optimal managed competition subsidies. *NBER Working Paper*, (25616), February 2019.

[74] J. Milyo and J. Waldfogel. The effect of price advertising on prices: Evidence in the wake of 44 liqormart. *The American Economic Review*, 89(5):1081–1096, December 1999.

[75] O.A. Morgan, J.C. Whitehead, W.L. Huth, and G.S. Martin. A split-sample revealed and stated preference demand model to examine homogenous subgroup consumer behavior responses to information and food safety technology treatments. *Working paper*, 2013.

[76] A. Nevo. Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2):307–342, March 2001.

[77] K. Nosal. Estimating switching costs in the market for medicare advantage plans. *Working paper*, 2012.

[78] D.J. Phaneuf, L.O. Taylor, and J.B. Braden. Combining revealed and stated preference data to estimate preferences for residential amenities: A gmm approach. *Land Economics*, 89(1):30–52, February 2013.

[79] M. Polyakova. Regulation of insurance with adverse selection and switching costs: Evidence from medicare part d. *American Economic Journal: Applied Economics*, 8(3):165–95, July 2016.

[80] R.O. Reid, P. Deb, B.L. Howell, and W.H. Shrank. Association between medicare advantage plan star ratings and enrollment. *The Journal of the American Medical Association*, 309(3):267–274, 2013.

[81] M. Rothschild and J. Stiglitz. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90(4):629–649, November 1976.

[82] M. Shum. Does advertising overcome brand loyalty? evidence from the breakfast-cereals market. *Journal of Economics and Management Strategy*, 13(2):241–272, 2004.

[83] K.A. Small and H.S. Rosen. Applied welfare economics with discrete choice models. *Econometrica*, 49(1):105–130, January 1981.

[84] J. So. Adverse selection, product variety, and welfare. evidence from medicare advantage. *jmp*, November 2019.

[85] Z. Song, M.B. Landrum, and M.E. Chernew. Competitive bidding in medicare advantage: Effect of benchmark changes on plan bids. *Journal of Health Economics*, 32(6):1301–1312, 2013.

[86] M. Spence. Monopoly, quality, and regulation. *The Bell Journal of Economics*, 6(2):417–429, Autumn 1975.

[87] G. J. Stigler. The economics of information. *The Journal of Political Economy*, 69(3):213–225, June 1961.

[88] J.E. Stiglitz. Imperfect information in the product market. *Handbook of Industrial Organization*, 1, 1989.

[89] L. Sun. Welfare impacts of supply-side regulation in medicare advantage market. *Working paper*, 2017.

[90] J. Swait and J. Louviere. The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research*, 30(3):305–314, August 1993.

[91] R. Town and S. Liu. The welfare impact of medicare hmos. *The RAND Journal of Economics*, 34(4):719–736, Winter 2003.

[92] K. Train. Welfare calculations in discrete choice models when anticipated and experienced attributes differ: A guide with examples. *The Journal of Choice Modelling*, 16(C):15–22, September 2015.

[93] K. Train and W.W. Wilson. Estimation on stated-preference experiments constructed from revealed-preference choices. *Transportation Research*, Part B(42):191–203, 2008.

[94] G.J. Wedig and M. Tai-Seale. The effect of report cards on consumer choice in health insurance market. *Journal of Health Economics*, 21:1031–1048, 2002.

[95] Y. Wu. Supply response to consumer inertia: Strategic pricing in medicare part d. *Working paper*, January 2016.

[96] L. Yang, C.F. Choudhury, and M. Ben-Akiva. Stated preference survey for new smart transport models and services: Design, pilot study and new revision. *Working paper*, 2009.

[97] J. Yeo and D.P. Miller. Estimating switching costs with market share data: An application to medicare part d. *International Journal of Industrial Organization*, 61:459–501, November 2018.

Vita

Alexandra Charbi was born in Athens, Greece on 20 March 1986, the daughter of Mr. Nikolaos Charbis and Mrs. Stella Charbi. She received her Bachelor of Science degree in Economics from Athens University of Economics and Business. Upon graduating she continued her studies in the same college where she earned a Masters in Applied Economics and Finance. She decided to continue her studies abroad and in 2012 she joined a Masters program in Economic theory at the University of Michigan in Ann Arbor and she graduated in 2014. Upon graduating she worked at a small economic consulting company based in Ann Arbor while applying for PhD studies. She was accepted by the University of Texas at Austin for enrollment in their Economics program and started her studies in August, 2015.

Permanent address: 1636K West 6th Str.
Austin, Texas 78703

This dissertation was typeset with LaTeX† by the author.

†LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.